Issue on tests and measures of positive adult development 2

BEHAVIORAL
DEVELOPMENT
BULLETIN

bao
JOURNALS

Volume **19** | Number **4** | December **2014**

BEHAVIORAL DEVELOPMENT BULLETIN

» MANUSCRIPT SUBMISSION

Please submit manuscripts electronically via email to the editors.
- Michael L. Commons: commons@tiac.net
- Martha Pelaez: martha.pelaez@fiu.edu

Guidelines for submission are available on the following webpage:
http://www.baojournal.com/BDB%20WEBSITE/Submissions.html

» THANKING OUR REVIEWERS

The work of reviewing is essential to the advancement of Adult Development theory and practice. We are grateful to the following individuals who reviewed for Behavioral Development Bulletin, issue on Tests and Measures of Positive Adult Development 2.

- Albert Erdynast, Ph.D., D.B.A. — Antioch University
- Anibal Gutierrez, Ph.D. — Florida International University
- Carl Oliver, Ph.D. — Loyola Marymount University
- Carolyn Adams-Price, Ph.D. — The National Strategic Planning & Analysis Research Center (NSPARC)
- Cory David Barker, MA — Antioch University
- Edward K. Morris, Ph.D.— University of Kansas
- Erik Jan van Rossum, Ph.D. — University of Twente
- Erik Mayville, Ph.D., BCBA-D — Institute for Educational Planning
- Gabriel Schnerch, M.A., Ph.D. Candidate, BCBA— University of Manitoba
- Hudson Golino, Ph.D. — The Federal University of Minas Gerais, Laboratory for Cognitive Architecture Mapping
- José Ferreira-Alves, Ph.D. — University of Minho
- Laurilyn D. Jones, Ph.D. Candidate — Oslo and Akershus University College, The Mechner Foundation
- Maria Judith Sucupira da Costa Lins, Ph.D. — Federal University of Rio de Janeiro, Brazil
- Michael D. Hixson, Ph.D., BCBA-D — Central Michigan University
- Michael F. Mascolo, Ph.D. — Merrimack College
- Patrice Marie Miller, Ph.D. — Salem State University, Harvard Medical School

» MISSION STATEMENT

Behavioral Development Bulletin (BDB) informs the field of developmental psychology by taking a behavioral analytic approach, including research in cognitive and emotional development, developmental theory, socialization, education, and speech-language pathology. It looks at the biological and environmental factors and systems that affect behavioral development, while maintaining primary interest in the role of reinforcement and environmental contingencies that influence behavior change.

BDB has three goals. The first is to understand human development in behavioral terms. The second is to reach out to developmental psychologists with the innovations that behavior-developmental approaches have provided. The third is to publish behavior analytic interventions that measure and promote development and change, specifically educational, clinical and organizational interventions that stimulate and increase the likelihood of development.

BDB publishes articles of an inter-and-multidisciplinary nature including areas of socio-biology, verbal behavior, and behavioral methodology. It is especially relevant to behavior analysts who study learning and the developmental processes responsible for behavior changes and their progressive organization.

The journal has five recurring special sections:

1. Behavioral Assessment and Intervention in Children

2. Developmental Approaches to Clinical and Counseling Psychology

3. Behavior-developmental Based Education

4. Positive Adult Development

5. Speech-Language Pathology and Applied Behavior Analysis

BDB is interested in articles that are intersections of behavioral methods and development. The articles are not to be interpretative, humanistic, or metaphorical. ■

Contents

Nancy Nordmann   Historically collected data and the analysis of that data are presented demonstrating the use of seminal measures of life satisfaction and interpersonal behavior in identifying a model of interpersonal functioning, autogenesis, the elaboration of which addresses an issue of contemporary focus, the integration of the interpersonal constructs of agency and communion. These seminal measures are demonstrated to be of continuing value as Index Measures of autogenesis and to provide criterion validation of the autogenetic model.

Autogenesis is proposed as an interpersonal dynamic defined as the development of the potential of a sense of origin, energy, self-direction, effectiveness and ultimately communion in one's activities as they relate to the world of others. Autogenesis is posited as expressing, through a number of thematic categories of interpersonal perception and functioning, the growth of the self as an ecologically adapted origin and agent. The suggested goal expressed by the autogenetic process is that of a self-directing, self-maintaining, self-realizing entity that is responsive in an interpersonally structured world. Hypotheses examined are: the reliability of identifying autogenetic stages in individual life interviews; the validity of identifying autogenetic stages using an Index Measure; and the criterion validity of the autogenetic model.

Preface

Michael Lamport Commons

Harvard Medical School

This issue of the *Behavioral Development Bulletin* will focus on the measurement of positive changes during adulthood. Adult development has multidimensional threads (Commons, 1999). These threads are sequences of events that run throughout the whole course of adult life. Most of previous work on measuring change in adulthood examined deficiency and decline. There is now a substantial body of evidence, however, that positive forms of development can occur at all periods of the lifespan, including adulthood. The four major forms of adult development are positive adult development, directionless change, stasis, and decline. The first of the four forms, positive adult developmental processes, is divided into at least six parts: hierarchical complexity (orders, stages), knowledge, experience, expertise, wisdom, and spirituality. Change, the second of the forms, is divided into periods, usually defined by decades and seasons, that concern the themes in the life course. One might think of it as developmental sociology. Over the last twenty years, there has been a rapid increase in measurement instruments that examine development during adulthood. Many of these measures are presented here in this issue. Many of the papers in this issue incorporate the *Model of Hierarchical Complexity* (mhc). The mhc is an axiomatic behavioral developmental model. It is an extension and revision of Inhelder and Piaget (1958) that also embodies previous theory of Pascual-Leone (1970) and Werner (1957). This model may be applied to study development in any domain, including adult development and development in any species. In addition, it also includes articles on stages and other models, which could be combined with mhc to develop better models for studying development.

The first section of this issue includes articles that present Mathematical Models of Constraints on Development. The opening article uses the *Model of Hierarchical Complexity* to examine how differences in rate of stage change results in a difference in the highest average stage (smarts") predicted to be attained by age 70. This paper shows how the rate of development predicts difference in stage attained at any age. Hence a stage theory provides a dynamic understanding of the accumulation of smarts over a lifetime. The paper does this by testing the relationship of stage of development to age. The second paper in this section examines if the stage of development an organism predicts the number of neurons it has. This paper hypothesizes that neural development is an evolutionary means to higher rate of reinforcement. The mechanism is that successfully performing more hierarchically complex tasks allows the attainment of more reinforcement. With more reinforcement, the chances of survival will increase. The next paper in the first section investigates the validity of the Cattell-Horn-Carroll (chc) model in one individual, using the intraindividual approach. The final paper in this section tests the effect of small variables on task performance. The task sequences used were from the Logic/mathematics/physical science subdomains. The four instruments used were the algebra, balance beam, infinity and laundry instruments. The variables tested were hierarchical complexity, place in order, the number of calculations needs, the size of the numbers, and the causal variable position.

The second section of this issue is a compilation of articles on theory, models and advances in the models. This section consists of papers that examine the long term debate and controversy on classification and numbering of stage domains. It also consists of papers that define the deconstructive pattern that emerges and reemerges during developmental or "stage" transition. This section also illustrates the evolution and advances in the mathematical and behavioral developmental models, such as the *Model of Hierarchical Complexity*.

The opening article of the second section is by Commons and Chen on the evolution and advances in the *Model of Hierarchical Complexity* (MHC). The paper traces the evolution of the MHC, within four periods. The two earliest periods were Commons' pre-college years, college and graduate school years, 1973 to 1982. After the Model was developed, there were advances in the period from 1982 to present. The paper also acknowledges the contributions of different people who played critical roles in the overall development, revisions and advances for the *Model of Hierarchical Complexity*. The next paper in this section uses the *Model of Hierarchical Complexity* (MHC) to test the theory that different domains in development would develop in synchrony, allowing an individual to solve tasks from various domains using the same mental structure for each task. The MHC instruments used were the empathy, helper person, counselor patient, breakup, caregiver, algebra, balance beam, infinity and laundry instruments. The third paper in this section uses a cross-sectional study with 180 adult subjects to research structural-developmental levels of conceptions of the beautiful. The paper also studied the relations between structural-developmental conceptions of the beautiful and structural-developmental conceptions of moral development. The final paper in this section defines the deconstructive pattern that emerges and reemerges during developmental or "stage" transition, and shows how it is necessary to incorporate this deconstructed "stage" into a more complex system of knowing or "stage".

The third section of this issue is a compilation of articles on the application of models and instruments for measuring development. Various authors describe the purpose, etiology, validity, and reliability of those models and instruments. They also explain the appropriate methodologies for their use in the domains of economics, eudaimonic wellbeing, education, leadership and adult development, culture and identity and others.

In this third section of the issue, the opening article summarizes findings of the ongoing interdisciplinary research project on differences in neural processing of individualistic vs. collectivistic oriented test persons (managers vs. non-managers). The paper also integrates dominant psychological trait theories (BIG-5) and theories of adult development (AD). The second paper in this section investigates the relationship between goal-pursuit (as a self-actualizing process), metacognition (as an attainable intellectual capacity), and (eudaimonic) well-being. The third paper in this section uses a measurement of cognitive bias to report American and Indian responses to environmental issues designed to trigger sunk cost bias. The paper explores how different factors influence reasoning biases, specifically, how people with individualist cultural values in the United States differ from people with collectivist cultural values in India. The fourth paper in this section compares learning performances of first-year and upper-class traditional-aged college students to determine if there were differences in learning for belief-consistent and belief-inconsistent information. The fifth paper in this section examines a particular group of midlife adults in "mixed-orientation marriages" who experience the re-emergence of previously denied motives. The paper describes the ways in which the inadequacies of available cultural scripts for sexual orientation impede their ability to re-integrate their same-sex desire into their adult identities. The final paper in this section reports an analysis identifying perspectives of interpersonal agency denoted as autogenesis. The suggestion is that historical data and analyses can be revisited as a source of relevant contemporary insights.

As can be seen, this issue presents a very rich and broad range of information on aspects of positive adult development. Putting this issue together would not have been possible without the help of managing editor, Sagun Giri. His efforts are very much appreciated.

**BEHAVIORAL
DEVELOPMENT
BULLETIN**

A model of stage change explains the average rate of stage of development and its relationship to the predicted average stage ("smarts")

Michael Lamport Commons[1], Leonard Sidney Miller[2], and Sagun Giri[3]

1 Harvard Medical School
2 University of California, Berkeley
3 Dare Institute

ABSTRACT

A number of different previous methods for measuring "smarts" have led to the model of hierarchical complexity (MHC), a context free neo-Piagetian mathematical model of behavioral complexity. It provides a way to classify tasks as to their hierarchical complexity. Using the model of hierarchical complexity, this study examines how differences in rate of stage change results in a difference in the highest average stage (smarts") attained by 70 year old adults. The average stage of development ("smarts") was shown to be predicted by the log of age with an $r = .79$. It uses data from Colby, Kohlberg, Gibbs, Lieberman (1983) to test the model. It also predicts that on the average there is one stage of development during adulthood.

KEYWORDS: stage limits, logage, age, IQ, smarts, smartness, adult stages

THERE HAS BEEN A long controversy about the relationship between "smarts" and biology and the environment. Most typically, the "smarts" being discussed in here has been measured using IQ. There are extensive reviews of this controversy so it will not be covered in this paper. Also there is an extensive literature on the relationship between stage and IQ (DeVries, 1974; Dudek, Lester, & Goldberg, 1969; Humphreys & Parsons 1979; Kohlberg & Devries, 1984; McClelland, 1973). In general, these studies have shown the relationship between stage and IQ to be modest. The problem with all of these studies is that only a small range of ages and stages have been used. This limitation would tend to attenuate the correlation between stage and IQ.

Among the problems with IQ as a measure of smarts, a major one is that there has been no psychophysical approach. IQ is based on a psychometric approach. Psychometrics depends on an analysis of only responses to items without any apriori theory about the meaning of the items or stimuli that lead to those responses. This kind of approach makes it difficult to interpret many results, because the nature of the items or stimuli, and the relationship between them, are not specified in advance. Psychophysics, on the other hand, depends on finding the relationship between in-

dependently scaled characteristics of stimuli and the responses to those stimuli. Consider that sound intensity is the scaled physical property of sound. The reported loudness is the response to the intensity of those sounds. One of the major problems with IQ is that we do not have an a-prior scaled difficulty of the IQ items. One could find out how difficult participants found the items based on their performance. In contrast to psychophysics, psychometrics was introduced when there was no understanding of how to scale the difficulty of items beforehand. As a result of the limitations of psychometrics in the case of IQ, it is not possible to figure out why the items in IQ tests have the difficulties they do. It is also not possible to explain why the items then "loaded" on the general factor. That is, why certain items are solved earlier and other items later can only be explained in a post hoc manner.

The problems with a number of different previous methods for measuring "smarts", but particularly the IQ measures, have led to the model of hierarchical complexity (MHC). The MHC is a psychophysical as opposed to a psychometric approach to the measure of "smarts." The order of hierarchical complexity is a property of the tasks that exists in the real world and is defined *a priori*. The order of hierarchical complexity (OHC) is an analytic measure applied to tasks. An organism's performance on those tasks is called *stage*. It should be noted here that this theory replaces earlier notions of stage (Inhelder & Piaget, 1958). Those notions, in a manner similar to IQ, confounded the stimulus and response

Correspondence regarding this article should be directed to Michael Lamport Commons, Ph.D., Assistant Clinical Professor, Department of Psychiatry, Harvard Medical School, Beth Israel Deaconess Medical Center, 234 Huron Avenue, Cambridge, MA 02138-1328, commons@tiac.net.

in assessing stage. The vast majority of stage-related data have been generated by simply scoring responses and ignoring the task or stimulus.

When stage is based on the *a priori* analytic scaling of items that is possible using the model of hierarchical complexity, stage does not have the problems of the psychometric method. Rather, as will be spelled out, the "Order of hierarchical complexity (OHC)" of a task predicts the performance on those tasks. The performance is measured by the Rasch scaled difficulty of the items (Commons, Goodheart, Pekker, Dawson, Draney, & Adams, 2008). Relating the order to the task performance in this way is a use of psychophysics not psychometrics.

There are several other advantages of the use of MHC to study "smarts." One important one is that the order of hierarchical complexity has been shown to be a property of tasks that is independent of the form, content and method of testing. It is not based on an analysis of performance but on an analysis of task demands. By using Rasch analysis of participant's performance, the obtained difficulty has been compared to the posited order of hierarchical complexity. The OHC predicts the Rasch scores with an *r* between .91 and .98, as shown in Giri, Commons and Harrigan (2014). This study also showed that factoring nine stage measures differing in form, content, and method yielded only one factor, stage. The implications of these results that the model of hierarchical complexity can be used with any kind of task from any domain.

A second important advantage is that the MHC includes stages of action and reasoning beyond the formal stage. Using the MHC to *a priori* and analytically determine the difficulty of the items on a typical IQ test (the WAIS-V), suggested that these items top out at the formal order. There were only a very small number of items characteristic of the next order, systematic (Chen, & Commons, 2014). Not only is the IQ test limited to having only a few items that require the first stage beyond the formal stage, it also does not go down below the primary stage. This was also determined from the above-mentioned coding of the items on the WAIS (Chen, & Commons, 2014). What is particularly problematic, since Commons (2008) estimated that about 20% of educated adults complete tasks that are more complex than the single variable problems that are required for the formal stage (Commons, Miller & Kuhn, 1982) is the relative lack of items that test for reasoning that is more complex than the formal order. The next order, systematic, involves successfully solving dilemmas or problems with two or more independent or "causal" variables.

The purpose of this paper is to see how the rate of development predicts stage attained at any age. The use of the model of hierarchical complexity allows for a dynamic understanding of the accumulation of "smarts" over a lifetime. The paper does this by

Table 1. order number and order name

order		moral maturity scores (MMS)
number	name	
0	computational	—
1	automatic	−150
2	sensory or motor	−100
3	circular sensory motor	−50
4	sensory-motor	0
5	nominal	50
6	sentential	100
7	preoperational	150
8	primary	200
9	concrete	250
10	abstract	300
11	formal	350
12	systematic	400
13	metasystematic	450
14	paradigmatic	500
15	crossparadigmatic	550
16	meta-crossparadigmatic	600

Note. adapted from "Correspondence between some life-span, stage" by Commons, M. L. & Tuladhar, C. T., 2014, *Behavioral Development Bulletin*, *19*(3), p. 26.

testing the relationship of stage of development to age. The longitudinal data set to be used was collected by Colby, Kohlberg, Gibbs and Lieberman (1983) using the moral maturity scale. The moral maturity scores were converted into hierarchical complexity stage scores (Tuladhar & Commons, 2014),

First, in order to allow for a better understanding of this material, the paper introduces the model of hierarchical complexity. It then presents a stage acquisition model of the factors necessary for development to take place. Third, a statistical model that shows the relationship between stage and maturation is presented. In that model, age represents maturation. Maturation is mostly due to biology but may be affected by environmental factors, including stimulation, nutrition and disease. What is to be examined here is how good age is at predicting stage is. Fourth, based on the relationship between stage and age, does stage progress at a diminishing rate with age?

» THE MODEL OF HIERARCHICAL COMPLEXITY

The model of hierarchical complexity (MHC) is a non-mentalistic, neo Piagetian mathematical model (Krantz, Luce, Suppes, & Tversky, 1971; Luce & Tukey, 1964). MHC deconstructs tasks into the actions that must be done at each order and this allows for the measurement of stage performance. MHC provides an analytic and *a priori* measurement of the difficulty of task actions and postulates that the difficulty is represented by the orders of hierarchical complexity (OHC) (Commons & Pekker, 2008). There are 17 known orders of hierarchical complexity. This is shown in Table 1, along with the corresponding Moral Maturity Scale scores (Colby & Kohlberg, 1987A, 1987B; Colby, Kohlberg, Gibbs, & Lieberman, 1983). Tuladhar and Commons (2014) have described the one to one relationship between moral maturity scores and MHC stages. Moral maturity scores are stages in the moral judgment subdomain.

Hierarchical complexity refers to the number of times that the coordinating actions must organize lower order actions. The hierarchical complexity of an action is determined by decomposing the action into the two or more simpler actions that make it up. This iterative process is done until the organization can only be carried out on a set of simple elements that are not built out of other actions. As shown in Figure 1, actions at a higher order of hierarchical complexity: *1)* are defined in terms of actions at the next lower order of hierarchical complexity; *2)* organize and transform the lower-order actions; *3)* produce organizations of lower-order actions that are new and not arbitrary, and cannot be accomplished by those lower-order actions alone. Once these conditions have been met, the higher-order action coordinates the actions of the next lower order.

To illustrate how lower actions get organized into more hierarchically complex actions, consider a simple example. Completing the entire operation $3 \times (4 + 1)$ constitutes a task requiring the distributive action. That distributive action non-arbitrarily orders adding and multiplying to coordinate them. The distributive action is therefore one order more hierarchically complex than the acts of adding and multiplying alone; it indicates the singular proper sequence of the simpler actions. Although simply adding can result in the same answer, people who can do both display a greater freedom of action. Thus, the order of complexity of the task is determined through analyzing the demands of each task by breaking it down into its constituent parts. For example, an order-three task can be broken down into a sequence of three concatenation operations. A task action of order three operates on two or more task actions of order two. A task action of order two operates on two or more task actions of order one.

Each task difficulty has an order of hierarchical complexity required to complete it correctly. Because tasks of a given order of hierarchical complexity require actions with the matching stage number to perform them, the stage of the participant's performance is equivalent to the order of complexity of the successfully completed task. Tasks are also quantal in nature. They are either completed correctly or not completed at all. There is no intermediate state. For this reason, the model characterizes all stages as hard and distinct. The quantal feature of tasks is thus particularly instrumental in stage assessment because the scores obtained for stages are likewise discrete.

In considering questions, such as the relationship between "smarts" and biology, or other questions related to "smarts", the model of hierarchical complexity presents several advantages. First, the three axioms (See Figure 1) make it possible for the model's application to meet real world requirements, including the empirical and analytic. In particular, earlier uses of arbitrary organizations of lower order of complexity actions, possible in the Piagetian theory, despite the hierarchical definition structure, leaves the functional correlates of the interrelationships of tasks of differential complexity ill-defined. Moreover, this model is consistent with the neo-Piagetian theories of cognitive development. According to these theories (e.g. Pascual Leone, 1970), progression to higher stages or levels of cognitive development is caused by increases in processing efficiency and working memory capacity. That is, higher-order stages place increasingly higher demands on these functions of information processing, so that their order of appearance reflects the information processing possibilities at successive ages. Finally, the MHC can be applied to any content and the behavior of any organism, not just humans.

A higher order action is:

1) defined in terms of the task actions from the **next lower order** of hierarchical complexity.

2) The higher order task action organizes two or more next lower order of hierarchical complexity.

3) The ordering of the lower task actions have to be carried out **non-arbitrarily**.
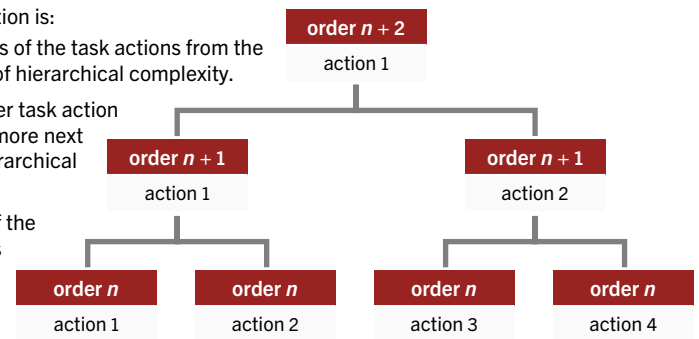
Figure 1. Shows the order of hierarchical complexity tree. Each higher order action organizes two or more next lower order actions. The hierarchical complexity of a task refers to the number of concatenation operations it contains, that is, what is the number of recursions that the coordinating actions must perform?

» THE RELATIVE ROLES OF MATURATION AND ENVIRONMENTAL CONTRIBUTIONS TO SMARTNESS ("SMARTS")

There are two main forms of stage change, which are developmental advances in behavior. Horizontal décalage refers to fact that once a certain task of a given order may be completed correctly in one task sequence, it is not necessarily initially completed correctly in another task sequence. In Baylor's (1975) words, "a horizontal décalage arises when a cognitive structure that can be successfully applied to task X cannot, though it is composed of the same organization of logical operations, be extended to task Y." One kind of stage change, therefore, is simply learning to apply a set of actions learned in one domain to another set of actions, at a similar order of hierarchical complexity, in another domain. Most developmental advances in behavior are of the horizontal form. In contrast, there is also real stage change, also sometimes called vertical décalage. This paper focuses on the vertical form.

From our perspective, there are two main factors necessary to developmentally advance behavior. They are conceptually separable but both are necessary. First is that there has to be a capacity to change. Second, the environment must be supportive of change.

1. There must be a "capacity" to change. This may be represented most compactly by Pascual-Leone's (1970) suggestion that to solve a problem at order N, there has to be a working memory of 2^N. This is what is termed capacity. This means that there are limits to what the environment produces in development at any age. Also the environment has to be tuned to the present performance of the organism to produce "maximal" stage change. No matter how much and how good the training is, at a given age there is an upper limit of stage that may be attained. Note that what can be trained at a given age varies across individuals. Capacity is assessed by finding where in a developmental sequence the organism is performing. There is ideal capacity. Ideal capacity is what would develop if the environment were perfectly tuned to maximize development. Capacity is also dynamic, changing with age and experience.

2. There must be contingent reinforcement for engaging in the task. There are two parts to this.

a) First, the task has to be appropriate for where the organism is performing in the sequence of tasks. For a task to be appropriate, it has to have some reasonable probability of being completed successfully. This requires first that the task is correctly placed in the developmental sequence. Second, the organism has to be functioning at that place in the sequence. One can overwhelm a student by giving tasks that are too much above the present stage at which they are functioning. One can bore a student by giving them work they already do perfectly.

b) Second, there must be some kind of reinforcing consequence that ensues from the completion of the task. The reinforcement is conceived of in much broader terms than what behavior analysts generally use. Reinforcement can include task mastery, which is set up by the drive of being interested, which is usually described as curiosity. Reinforcement may also be social recognition and attention.

The argument to be presented here does not suggest that there is not a contribution of environment to "smarts". With reinforcement for correct answers in the laundry problem, 5th and 6th graders moved from 100% reasoning at the concrete stage to 75% reasoning at the formal stage (Commons, Davidson & Grotzer, 2007). Our preliminary data on training with reinforcements in a non-literate Nepal sample shows a great increase in formal stage answering and even some systematic stage answering (Commons, Giri, & Tuladhar, 2013). These studies show that environment did contribute to smarts. But for the environment to contribute to smarts there has to be reinforced training. There also has to be transfer of training tests to exclude the possibility of rote learning.

» ENVIRONMENTALLY EFFICIENT METHOD OF MAKING CHANGES IN BEHAVIOR STAGE

Based on the above discussion, we propose the following formal model of the factors necessary for behavioral stage change. The change in behavioral stage, Δ_B, is simply the product of the time actively engaged in getting the right answers to a task when placed in the developmental sequence correctly. This is shown in the following equation:

$$\Delta_B = t \times pl, \qquad (1)$$

where Δ_B is the change in behavior, $t = f_1(S_{\text{Contingencies of reinforcement for correct answers}})$, and $pl = f_2$(being placed in the right place in the developmental sequence).

The more time spent alive inevitably leads to more information being processed. More information being processed leads to higher chances of coordinating the information to higher stages. Time engaging actively on a task is sensitive to contingent reinforcement of correct responses.

Consider the probability of a 7-year old effectively engaged in the task of solving the equation $x - 1 = 0$. Most children at 7 years of age generally perform at the primary stage 8. This child would

also most likely be performing at the primary stage, two stages before the formal stage 11. Because this is a formal order 11 task, the child will likely fail it. So a task being placed too high in a task sequence leads to failure if the participant is not performing at the same stage and substage in the developmental sequence. Similarly, a task being placed too low on the task sequence leads to boredom on the part of the participant, and a failure to engage in the task. Because the person already does the task, learning does not occur as well.

» MEASURING THE AVERAGE STAGE OF DEVELOPMENT (SMARTS) USING MHC AS A FUNCTION OF AGE

As argued above, behavioral stage as measured by the model of hierarchical complexity gives a relatively unbiased measure of what we call "smarts." After establishing the usefulness of stage as a measure of "smarts", this paper presents a general and simple notion, that what produces behavior change is the amount of time spent engaged with environmental tasks presented at an appropriate order of complexity and accompanied by effective reinforcing events. This is the Δ_B equation discussed above. This is a very general notion. Because it is based on the MHC, it also is not based on specific context or content or any specific intervention. It can apply to any organism at any order of development. Even though there is evidence that measured behavioral stage may be increased over time based on experience and maturation, there still may be limits as to how far training can get. Therefore a second and separate question that can be pursued here is what are the constraints to development at a given age?

The next part of this paper uses some empirical data to see what the best predictor of "smarts" or the organism's currently measured behavioral stage is. We start with the simplest model, which examines the relationship between stage ("smarts") and age (maturation). The explanation for the relationship between stage ("smarts") and age (the maturational contribution) is simple. Individuals develop as long as they are appropriately stimulated and supported by the environment, as shown in the Δ_B equation discussed above. They also have to have enough capacity to engage with increasingly hierarchically complex tasks. As long as more complex problems and dilemmas are presented by the environment, there will be an increase in stage under these conditions. In other words, consider age, stage, and the order of hierarchical of complexity of the tasks presented, the model of hierarchical complexity provides an explanation for how stage change results in average stage attained at a given age.

Four steps show what determines the stage attained

The steps of the derivation of the regression equation for predicting the average stage attained by the participants as a function of age is shown next.

Step 1: recall that there are the differences between the order of hierarchical complexity of tasks and the corresponding stage of performance on those tasks. Order of hierarchical complexity (OHC) is an *a priori* analytic measure of difficulty applied to tasks. Stage is a performance measure of the most hierarchically complex task solved by the organism in question.

Step 2: Based on the information in Step 1, it is important to start out by deriving a measure of stage that can be predicted. This starts with the following definition.

Table 2. mean and *SD* for the variables

mean (*SD*) MMS	mean (*SD*) stage (MHC)	mean (*SD*) age	mean $\log_2$ age (*SD*)
M = 306.78; *SD* = 71.47	*M* = 10.13; *SD* = 1.15	*M* = 20.9; *SD* = 7.51	*M* = 1.27; *SD* = .19

Total amount of hierarchical complexity of a task $= 2^N_{OHC}$.

This is based on Pascual-Leone's (1970) suggestion that to solve a problem at order *N*, there needs to be a working memory of 2^N. Due to the definitions of stage and of order given in Step 1, the stage number, N_{stage}, is the same as the number N_{OHC}, for the most hierarchically complex task solved

$$2^N_{OHC} = 2^N_{stage}.$$

Therefore, what we will be predicting is N_{stage} or performance. Individual scores will be predicted. The predictions for individual scores will include all valid data for all participants.

Step 3: Determine how age will be considered. The prediction is that *Age* = *t* helps determine the amount of hierarchically complex information processed correctly. As asserted in Step 1, stage is the order of hierarchical complexity of a task (OHC) completed correctly. But using age in terms of simple number of years would not be appropriate as explained next. It is asserted and to be tested that $N_{stage} = \log_2(t)$. This equation shows that the more time spent alive inevitably leads to more information being processed correctly. What is to be predicted is N_{stage}, which is

$$\log_2\left(2^N_{stage}\right), N_{stage} = f_0\left(\log_2\left(2^N_{stage}\right)\right).$$

Because we take the $\log_2(2^N)$ to get stage, we have to take the $\log_2$ of age: $\log_2(age)$. We only explore one predictor, $\log_2(age)$ to see whether it is a good predictor of behavioral stage.

Step 4: The more information being processed at a given time leads to higher chances of coordinating the information at higher orders of hierarchical complexity. Therefore, at the fourth step, it will be shown that development is set by the rate of stage change, the parameter *κ*. The parameter *κ*, by definition, is the rate of change of stage with age. The rate of change is represented as a partial derivative noted as ∂. The derivative of stage ∂*N*/∂*t* with respect to time is

$$\frac{\partial}{\partial t}\left(\log_2\left(2^N\right)\right) = K.$$

Then N_{stage} is substituted for $\log_2(2^N)$ yielding

$$\frac{\partial}{\partial t}\left(N_{stage}\right) = K.$$

By using N_{stage} as a stage variable, *κ* may be found for individuals or for groups of individuals. In sum, this is:

$$K = \frac{\partial}{\partial t}\left(N_{stage}\right),$$

which is the partial derivative of N_{stage} with respect to *t*. Because

N_{stage} is a function of age, *κ*, which may be found for an individual or for groups, requires the age of the individual or the average age of the group.

» MORAL MATURITY SCORE WAS REGRESSED ON $\log_2$(AGE) OF PARTICIPANTS

In this paper we look at individual Moral Maturity Scores and the corresponding stage from the model of hierarchical complexity. We will argue that when one looks at individuals, there are two contributions to Moral Maturity Scores: *a)* the maturational and *b)* the environmental. Training, as shown by Δ*B* is one of the contributors to individual differences and behavioral change up until the limits imposed by maturation.

The data used in this analysis come from Colby, Kohlberg, Gibbs, Lieberman (1983). They presented three longitudinal studies on Moral Judgment. They reported individual Moral Maturity Scores. The data set consists of 51 people with multiple observations per person of their moral maturity assessed at approximately four year intervals. There were 225 observations in total. Some participants are missing some assessments.

» STATISTICAL ANALYSIS

The purpose of this section is to explain the statistical analysis employed to estimate how moral maturity is related to the logarithm, base 2 of the age of person at time of assessment.

First, to make clear that moral maturity scores (MMS) have an equivalent stage of development as the model of hierarchical complexity stage score, we show the simple linear relationship between the two (Tuladhar, & Commons, 2014). The following simple equation represents that result.

$$\text{Behavioral developmental stage} = 4 + 0.02\text{MMS}$$

Second, the moral maturity at an age of assessment is specified as linearly related to the logarithm, base 2, of age of the individual at assessment, as in Equation 2. The equation to be fitted is Equation 2

$$\text{MMS}(t) = \beta_0 + \beta_1 \log_2(\text{age}) + \varepsilon(t), \quad (2)$$

where *t* is the age of an observation at time of assessment, $\log_2(age)$ denotes the base 2 logarithm of the age of an individual at assessment, β_0 denotes the scale constant relating the offset between Moral Maturity Scores and $\log_2(age)$, and β_1 denotes the slope. It is the slope that we get from the regression between MMS and $\log_2(age)$. Additionally, $\varepsilon(t)$ is an error term, whose expected value is zero and whose variance is σ_ε^2, and ρ (Rho Greek letter) denotes the correlation between two temporally adjacent test scores for the same person.

The dependent variable MMS(*t*) denotes an interval level measure of moral maturity at the assessment that occurred at age *t*.

See the Appendix A for definitions of terms, and a more detailed explanation of the statistical model.

» RESULTS

Figure 2 shows the scatter plot of the data. Full information maximum likelihood estimates of the parameters of Equation (2) account for the effects of the repeated measurements and missing measures (See Appendix A). Table 3 (which follows below) presents the maximum likelihood parameter estimates and their statistical significance. Average moral maturity scores were predicted by $\log_2$(age): $r(225) = 0.785$; the adjusted $R^2 = 0.6163$, which is the proportion of the variance explained by the model; $p = .00014$ and the non-adjusted $R^2 = 0.6214$.
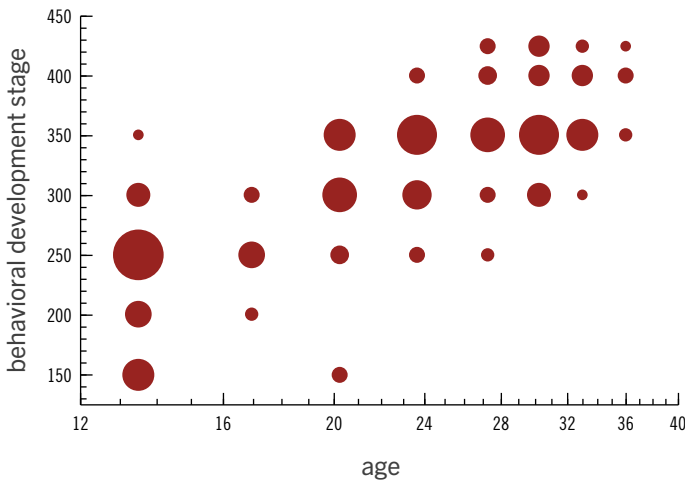
Figure 2. Shows the scatter plot for individuals. The age of participants was plotted on the *x*-axis ($\log_2$ scaled) and behavioral developmental stage was plotted on the *y*-axis. The size of the circles represents the number of cases.

From above and Table 3,

$$\text{MMS} = -73.92 + 88.17 \log_2 (\text{age}) + \varepsilon (t), \qquad (7)$$

where the expected value of $\varepsilon(t) = 0$.
From above, we have the equation showing the simple linear relationship between moral maturity scores (MMS) and stages of development,

$$\text{Behavioral development stage} (t) = 4 + 0.02 \text{MMS} (t). \quad (8)$$

Substituting the MMS into that equation yields

$$
\begin{aligned}
&= 4 + 0.02 \left(-73.92 + 88.17 \log_2 (\text{age}) \right) \\
&= 4 - 1.48 + 1.76 \log_2 (\text{age}) \qquad (8) \\
&= 2.52 + 1.76 \log_2 (\text{age}).
\end{aligned}
$$

Next, the values of κ are derived. As said previously κ is the rate of change of stage with age. From Appendix C, the estimated

value of κ was found for $\delta(N_{\text{stage}})/\delta$, the partial derivative of N_{stage} with respect to t.

Hence,
$$
\begin{aligned}
K &= \frac{\partial}{\partial t} (N_{stage}) \quad \text{(see Appendix C)} \\
&= \frac{2.539}{\partial (\text{age})}.
\end{aligned}
$$

» RELATIONSHIP BETWEEN STAGE AND AGE

In this section, it will be shown how small changes in β_1 result in large changes in stage, given age. This is because the long term cumulated effect changes the ultimate stage a person may reach. Note, that at any time, the predicted average stage is increasing. This is roughly a somewhat similar argument to IQ and the rate of learning. But it differs in that there is no arbitrary end, due to the measurement instrument, as seen with IQ. That is because MHC tests start in infancy and go through to the end of adulthood. The laundry task, for example, starts in the circular sensory-motor stage 3 (age .8) and goes through to metasystematic stage 13.

Using Equation 9, the predicted stages across ages are shown in Figure 3, given different changes in values assumed for the value of β_1. The first change is to make a 10% increase in β_1 and then a 20% increase in β_1. Generally, each stage is one standard deviation from the next one. In IQ, 15 points is one standard deviation. Depending on age, 10% will predict a different standard deviation for stage as seen in Figure 3. The second change is to make a 10% decrease in β_1 and then a 20% decrease in β_1. See appendix B for the mathematics underlying these results. What Figure 3 shows is that as people get older, the maximum stage is never attained. That is, the slope of the lines flattens out as one ages.

Legend:
- $4 + 0.02(\beta_0 + 1.2\beta_1 \log_2(\text{age}))$
- $4 + 0.02(\beta_0 + 1.1\beta_1 \log_2(\text{age}))$
- $4 + 0.02(\beta_0 + \beta_1 \log_2(\text{age}))$
- $4 + 0.02(\beta_0 + 0.9\beta_1 \log_2(\text{age}))$
- $4 + 0.02(\beta_0 + 0.8\beta_1 \log_2(\text{age}))$
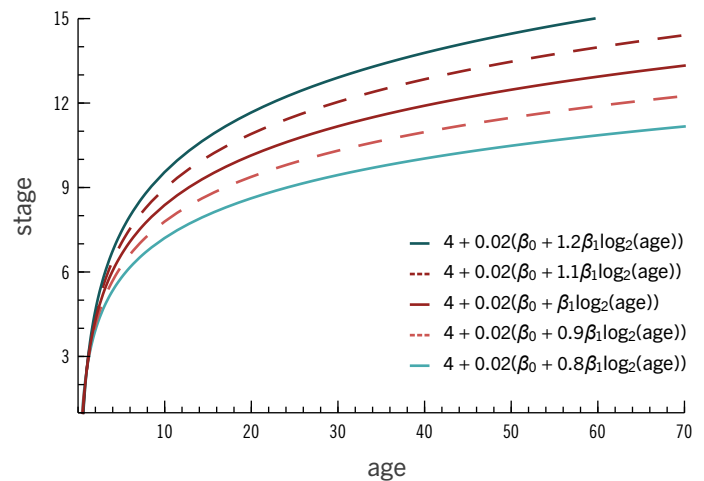
Figure 3. the relationship between age and stage

Table 3. model parameter estimates and statistical significance

parameter	parameter estimate	standard error of estimate	t-value	p-value
β_0	-73.92	19.03	-3.88	0.00014
β_1	88.17	4.41	20.00	1.79×10^{-51}
σ_v^*	39.09	1.91	20.52	0.0
ρ	0.46	0.07	6.91	4.99×10^{-11}

*Note. * represents 1 tailed test*

This 'slowing down' in development is shown in terms of specific predicted stages for each decade in adulthood in Table 4. While the average change in predicted stage across these 5 decades is .63, it can be seen that stage change in the decade from age 20 to age 30 is predicted to be twice what it is in the decade between 60 and 70.

Table 4. predicted stages for different adult ages (for the average case)

age	predicted stage	change in predicted stage
20	10.13	—
30	11.16	1.03
40	11.89	0.73
50	12.45	0.56
60	12.92	0.47
70	13.31	0.39

Table 2 showed that at the average age of the participants, which is 20.9 years, the average stage is around the Abstract Stage, 10.1. Table 5 shows that at the age of 8, the average stage is 7.6 (transitional to Primary 8). On the other hand, if one is in the top 20 percent, the stage of 8.8 is already transitional to concrete, a standard deviation above the mean. Note also that this small difference at age 8 becomes a large difference at age 70. If one is in the top 20 percent, one reaches the Crossparadigmatic stage 15, which is an overestimation from the model. This is because there is over extrapolation from the data. If one is down 20 percent, one only reaches the abstract stage 10, which makes sense.

Table 5. stage equivalent at age 8 and age 70 with 10 and 20 percent increase in β_1 and 10 and 20 percent decrease in β_1

age	average stage equivalence	stage with 10% increase in β_1	stage with 20% increase in β_1	stage with 10% decrease in β_1	stage with 20% decrease in β_1
8	7.6	8.2	8.8	6.2	6.8
70	13.1	14.2	15.1	11.9	10.9

» RELATIONSHIP BETWEEN CHANGE IN STAGE AGE DUE TO CHANGE IN AGE

This section will show the relationship between changes in stage due to changes in age (see Figure 4) under the same conditions as above: First the relationship is shown with a 10 percent increase in the estimated β_1 and then a 20 percent increase in the estimated β_1. Second, the relationship between age and the change in stage is shown due an increase in age with a 10 percent decrease in the estimated β_1 and then a 20 percent decrease in the estimated β_1. See Appendix C for the math and the model.

The dynamic representation of change in stage changes at different points in the lifespan. Essentially the rate of change decreases with age. What is somewhat startling is how much a difference the estimated slope coefficient makes on that rate of change.
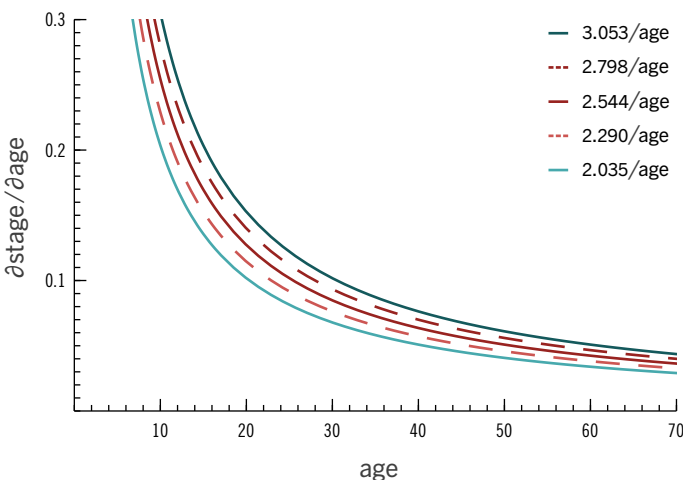
Figure 4. This plot shows the relationship between age and change in stage. These expressions are the derivatives of the expressions in Figure 3 with respect to age.

» DISCUSSION

We have argued that the idea that "smarts" can be measured with behavioral stage. Stage is a better measure than something like IQ because with the model of hierarchical complexity, the *a priori* difficulty of the items is known. This makes it possible to understand what the results mean in terms of what exactly was being tested.

Second, using notions from MHC, one can then construct powerful models of what predicts development. Specifically, this study presented an empirical test of a mathematical model of the average attained stage of development ("smarts"). It uses data from Colby, Kohlberg, Gibbs, Lieberman (1983) to test the model.

The first major result of this approach is the demonstration that stage increases as function of $\log_2(age)$. The correlation between age and stage was .79, accounting for a surprisingly large amount of the variance in stage (61%).

A second major result is that as people get older, one can see that a maximum stage is never attained. Here we proposed that based on a single parameter, κ, MHC provides an explanation for how differences in rate of stage change result in a difference in average stage, given age. This parameter, κ, is equal to the derivative of moral maturity stage of a person's performance with respect to time. It is suggested that κ is a measurement of increase in "smarts" as age increases. As age increases, the value of κ decreases and vice versa. What is constant throughout the life span as far as our limited data shows is the product of κ and age, 2.539. This finding is because the increase in stage is less and less with increasing age.

A third major result was the finding that small changes in β_1 (slope from the regression between stage and $\log_2(age)$) results in an increasingly large change in stage. This is because the long-term cumulated effect changes the ultimate stage a person may reach.

A fourth major result is that, on the average, there is less than one stage of development predicted during each decade of adulthood.

Each of these findings has important implications for the study of development, and particularly in the areas of aging and positive adult development.

Stage of development is the best measure of "smarts" and should replace IQ for individuals (Commons & Ross, 2008). It only loads on a single factor no matter the content or context of the instrument used to measure it, as shown by Giri, Commons, & Harrigan (2014). Because age is such a strong predictor of stage, people should consider the maturational contribution to stage of development.

Stage measurement has problems however. It is sensitive to a number of things. These include the reinforcement history for giving the highest stage performances on tasks. Many traditional cultures either punish high stage responses or at least do not reinforce them (Commons, Galaz-Fontes &Morse, 2006; Commons, Giri & Tuladhar, 2014; Day, 2008). This is the case in all subdomains even though we find only find one overall domain; stage (Giri, Commons, & Harrigan, 2014). For example in Ravnican's (2013) study, people gave traditional and concrete stage 9 answers to the version of the laundry problem she used. Day (2008) found

that religious fundamentalists gave concrete stage 9 answers to the saying "he who is without sin, cast the first stone" and other moral dilemmas. This extended to other moral stage answers. By using training, one can overcome these limitations, as we are showing in a study in Nepal (Commons et al., 2013). The main solution is to use simpler stage tasks and reinforce correct answers and then test with a similar set of tasks.

The importance of this paper is twofold. First, this paper addresses the issue of talking about individual differences of smartness with one parameter. This parameter is κ, the rate of change of stage with Age. The behavioral sciences have tried using IQ tried to explain in-dividual differences but IQ does not explain this well because there was no *a priori* idea of what the difficulty of the IQ items were. This made it extremely difficult to interpret what exactly IQ measured and therefore what the individual differences could be attributed to. Second, stage has a clear notion that attempted task solutions are different at each stage, with the higher stage action building upon the lower stage actions. There is an underlying logic to the progression of answers in the stage sequence. Stage is the equivalent to mental age. Stage divided by age is the equivalent to IQ, with the difference that it can be applied to tasks across the lifespan. By combining stage with age, one derives the average potential to learn. ∎

REFERENCES

Arthur, S., Goldberger, A. S. & Manskim, C. F. (1995). The Bell Curve by Herrnstein and Murray, *Journal of Economic Literature*, *36*(2), 762–76.

Baylor, G. W. (1975). Experiments in seriation with children: Towards an information processing explanation of the horizontal décalage, *Canadian Journal of Behavioral Science*, *7*(1), 29-29. (ISSN: 0008-400X).

Chen, S. J. & Commons, M. (2014, July). Order of Hierarchical Complexity Scores for the WAIS-IV Verbal Subtests, Presented at *European Society for Research in Adult Development*, Helsinki, FI.

Colby, A. & Kohlberg, L. (1987a). *The measurement of moral judgment: Vol. 1. Theoretical foundations and research validation*. New York: Cambridge University Press.

Colby, A. & Kohlberg, L. (1987b). *The measurement of moral judgment: Vol. 2. Standard form scoring manuals*. New York: Cambridge University Press.

Colby, A., Kohlberg, L., Gibbs, J., & Lieberman, M. (1983). *A longitudinal study of moral judgment*. Monographs of the Society for Research in Child Development, *48*(1-2). Chicago: University of Chicago Press.

Commons, M. L. (2008). Introduction to the model of hierarchical complexity and it relationship to postformal action. *World Futures: Journal of General Evolution* 65(1-3), 305-322.

Commons, M. L. Galaz-Fontes, J. F. & Morse, S. J. (2006). Leadership, Cross-Cultural Contact, Socio-Economic Status, and Formal Operational Reasoning about Moral Dilemmas among Mexican Non-Literate Adults and High School Students. *Journal of Moral Education*, *35*, 247-267.

Commons, M. L., Giri, S. & Tuladhar, C. T. (2013, June). An evolutionary-behavioral-developmental theory of stage development based on knowledge of tool usage. Presented at *Society for Research in Adult Development*, Salem, MA.

Commons, M. L., Goodheart, E. A., & Bresette, L. M. with Bauer, N. F, Farrell, E. W., McCarthy, K. G., Danaher, D. L., Richards, F. A., Ellis, J. B., O'Brien, A. M., Rodriguez, J. A. & Schrader, D. (1995). Formal, systematic, and metasystematic operations with a balance-beam task series: A reply to Kallio's claim of no distinct systematic stage. *Journal of Adult Development, 2*(3), 193-199.

Commons, M. L., Goodheart, E. A., Pekker, A., Dawson, T. L., Draney, K. & Adams, K. M. (2008). Understanding Rasch measurement: Using Rasch scaled stage scores to validate orders of hierarchical complexity of balance beam task sequences. *Journal of Applied Measurement, 9*(2), 182–99.

Commons, M. l., Grotzer, T. A. & Davidson, M. N. (in press). The necessity of reinforcing problem solutions for transition to formal operations: An examination of Piaget's equilibrium theory of stage change.

Commons, M. L., Li, E. Y., Richardson, A. M., Gane-McCalla, R., Barker, C. D. & Tuladhar C. T. (2014). Does the Model of Hierarchical Complexity produce significant gaps between orders and are the orders equally spaced? *Journal of Applied Measurement*. 15(4), 422-450.

Commons, M. L., Miller, P. M., & Kuhn, D. (1982). The relation between formal operational reasoning and academic course selection and performance among college freshmen and sophomores. *Journal of Applied Developmental Psychology, 3*, 1-10.

Commons, M. L. & Pekker, A. (2008). Presenting the formal theory of hierarchical complexity. *World Futures: Journal of General Evolution*, *64*(5-7), 375-382. Doi: 10.1080/02604020802301204

Commons, M. L. & Ross, S. N. (2008). Using the universality of hierarchical complexity: toward a cross-species measure of general intelligence. *World Futures: Journal of General Evolution.* 65(1-3), 383-398.

Commons-Miller, L. A. H. & Commons, M. L. (2007). Speciation of superions from humans: is species cleansing the ultimate form of terror and genocide? *Journal of Adult Development, 14*(3-4), 122-125.

Day, J. (2008). Conscience: Does religion matter? Empirical studies of religious elements in pro-social behaviour, prejudice, empathy development, and moral decision-making. In *The structure and development of conscience,* ed. W. Koops, D. Bruggman, and A. Sander, Eds. London: Psychology Press.

DeVries, R. (1974). Relationships among Piagetian, IQ, and achievement assessments. *Child Development, 45*, 746-756.

Dudek, S. Z., Lester, E. P., & Goldberg, J. S. (1969). Relationship of Piaget measures to standard intelligence and motor scales. *Percepinal and Motor Skills, 28*, 351-362.

Giri, S., Commons, M. L. & Harrigan, W. J. (in press). There is Only One Stage Domain. *Behavioral Development Bulletin*.

Humphreys, L. G. & Parsons, C. K. (1979). Piagetian Tasks Measure Intelligence and Intelligence Test Assess Cognitive Development: A Reanalysis. *Intelligence 3*, 369-382

Jensen, A. R. (1980). *Bias in Mental Testing*. New York: Free Press.

Jensen, A. R. (2002). Psychometric g: Definition and substantiation. In R. J. Sternberg & E. L. Grigorenko (Eds.). *The general factor of intelligence: How general is it?* (pp. 39–53). Mahwah, NJ: Lawrence Erlbaum.

Jones, H. E. & Bayley, N. (1941). The Berkeley Growth Study. *Child Development*, *12*, 167-173.

Judge, G. G., Hill R. C., Griffiths, W. E. Lukepohl, H. & Lee, T-C. (1988). *Introduction to the Theory and Practice of Econometrics* (2nd Eds.). Massachusetts: John Wiley & Sons.

Krantz, D. H., Luce, R. D., Suppes, P. & Tversky, A. (1971). *Foundations of measurement*, Vol. I: Additive and polynomial representations. (pp. 93-112). New York: Academic Press.

Kohlberg, L. & Devries, R. (1984). Psychometric and Piagetian intelligences: toward resolution of controversy. *Intelligence, 8*, 67-91.

Kuhn, D., Langer, J., Kohlberg, L. & Haan, N. S. (1977), The development of formal operations in logical and moral judgment. *Genetic Psychology Monographs*, *95*, 97–188.

Luce, R. D. & Tukey, J. W. (1964). Simultaneous conjoint measurement: a new scale type of fundamental measurement. *Journal of Mathematical Psychology*, *1*(1), 1–27. Doi: 10.1016/0022-2496(64)90015-X

Lynn, R. (1999). The attack on the bell curve. *Personality and Individual Differences*, *26*, 761–65.

McClelland, D. C., (1973) Testing for competence rather than for "intelligence". *American Psychologist*, *28*(1)1-14.

Moffitt, T. E., Gabrielli, W. F., Mednick, S. A., & Schulsinger, E (1981). Socioeconomic status, IQ, and delinquency. *Journal of Abnormal Psychology*, *90*, 152-156.

Neisser, U. Boodoo, G. Bouchard, T., Boykin, A. Wade; Brody, Nathan; Ceci, Stephen J.; Halpern, Diane F.; Loehlin, John C.; Perloff, Robert; Sternberg, Robert J.; Urbina, Susana (1996). "Intelligence: Knowns and Unknowns" *American Psychologist 51*(2), 77–101.

Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica, 32*, 301-345.

Ravnican, S. (2013). Employee Assessment using the Model of Hierarchical Complexity (MHC): Case Study (Master dissertation).

Rose, C. & Murray, D. S. (2011). *Mathematical Statistics with Mathematica*. Massachusetts: Springer.

Stokes, B. (2011). Mathstatica 2.5. *Journal of Statistical Software*, *47*(1).

Thisted, R. A. (1988). *Elements of Statistical Computing*. London: Chapman and Hall.

Tittle, C. R. & Rotolo, T. (2000). IQ and Stratification: An Empirical Evaluation of Herrnstein and Murray's Social Change Argument. *Social Forces*, *79*(1), 1–28.

Tuladhar, C. T. & Commons, M. L. (2014). Correspondence between some lifespan stage theory developmental sequences of stages and levels. *Behavioral Development Bulletin*, *19*(3), 24-27.

APPENDIX

» APPENDIX A
Parameters estimates

It is postulated that the errors arise from a first-order autoregressive process known as the AR-1 model. An autoregressive (AR) model is a representation of a type of random process. As such, it describes certain time-varying processes in nature, economics, etc. The autoregressive model specifies that the error depends linearly on its own previous values. In the first autoregressive model, AR-1 model, the error in equation [2] at time t is correlated with the error at time $t - 1$, as in equation [3],

$\varepsilon(t) = \rho \varepsilon[t - 1] + v(t)$ [3]

Where,

$\varepsilon(t)$ = the error at time t

ρ (Rho Greek letter) denotes the correlation between two temporally adjacent test scores for the same person.

$v(t)$ is an error term whose expected value is zero and whose variance is σ_v^2

If ε is a column vector of errors, $\{\varepsilon[1], \varepsilon[2], ..., \varepsilon[n]\}^T$, then the covariance of the errors in the model, $\Phi = E[\varepsilon \varepsilon']$, for a person with three observations, for example, has the form given by equation [4]. Note, ε is a column vector of errors but is written as a row vector, with T indicating that this row vector is transposed.

$$\varphi = E[\varepsilon \varepsilon'] = \frac{\varphi_v^2}{1 - \rho^2} \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix} \qquad (4)$$

Where,

ρ (Rho Greek letter) denotes the correlation between two temporally adjacent test scores for the same person

$\Phi_v{}^2$ denotes the variance of $v(t)$. It denotes the common variance held in all the observations and all the errors.

Capital σ_v has a matrix form.

Φ denotes the covariance of the errors

$\Phi = E[\varepsilon \varepsilon']$ is the model of errors for a person with three observations

Assuming the $v(t)$ errors have normal distributions, the collection of ε errors for an observation will have a Multinormal distribution. Hence, to continue the example of three observations for a person, the probability of the mm(t) and $\log_2$(age) observations are given by equation [5], and how the various parts of it are calculated.

$$\text{PDF}\left[\text{MultinormalDistribution}\left[\{0, 0, 0\}, \frac{\sigma_v^2}{1 - \rho^2} \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix}\right]\right] \quad (5)$$

$mm[1] - (\beta_0 + \beta_1 \log_2(t[1]))$,
$mm[2] - (\beta_0 + \beta_1 \log_2(t[2]))$,
$mm[3] - (\beta_0 + \beta_1 \log_2(t[3]))$

Where,

t = the age of an observation at time of assessment

$t[1]$ = the age of an observation at time 1 of assessment

$t[2]$ = the age of an observation at time 2 of assessment

$t[3]$ = the age of an observation at time 3 of assessment

$\log_2$(age) denotes the logarithm, base 2, of the age of an individual at assessment.

β_0 denotes the scale constant relating the offset between Moral Maturity Scores and $\log_2$(age)

β_1 denotes the slope. It is the slope that we get from the regression between MMS and $\log_2$(age).

σ_v^2 denotes the variance of $v(t)$. See equation 3. It denotes the common variance held in all the observations and all the errors.

$$\frac{1}{2\sqrt{2}\pi^{\frac{3}{2}}\sqrt{\frac{(-1 + 2\rho^2 - \rho^4)\sigma_\gamma^6}{(-1 + \rho^2)^3}}}$$

$$= \frac{1}{2}\left(-\frac{\left(-\beta_0 - \frac{\beta_1 \text{Log}[t[1]]}{\text{Log}2} + mm[1]\right)\left(-\beta_0 \text{Log}2 + \beta_0 \rho \text{Log}2 - \beta \text{Log}[t[1]] + \beta_1 \rho \text{Log}[t[2]] + \text{Log}2\, mm[1] - \rho \text{Log}2\, mm[2]\right)}{\text{Log}2\, \sigma_\gamma^2} \right.$$

$$- \frac{\left(-\beta_0 - \frac{\beta_1 \text{Log}[t[2]]}{\text{Log}2} + mm[2]\right) \times \left(-\beta_0 \text{Log}2 + 2\beta_0 \rho \text{Log}2 - \beta_0 \rho^2 \text{Log}2 + \beta_1 \rho \text{Log}[t[1]] - \beta_1 \text{Log}[t[2]] - \beta_1 \rho^2 \text{Log}[t[2]]\right)}{\text{Log}2\, \sigma_\gamma^2}$$

$$+\beta_1 \rho \log[t[3]] - \rho \log_2 mm[1] + \log_2 mm[2] + \rho^2 \log_2 mm[2] - \rho \log_2 mm[3])$$
$$\overline{\log_2 \sigma_\gamma^2}$$

$$-\left(-\beta_0 - \frac{\beta_1 \log[t[3]]}{\log_2} + mm[3]\right)\left(\left(\frac{\rho^2\left(-\beta_0 - \frac{\beta_1 \log[t[1]]}{\log_2} + mm[1]\right)\sigma_\gamma^2}{1-\rho^2} + \frac{\left(-\beta_0 - \frac{\beta_1 \log[t[3]]}{\log_2} + mm[1]\right)\sigma_\gamma^2}{1-\rho^2}\right)\right)$$
$$\overline{\frac{\sigma_\gamma^8}{(1-\rho^2)^4} - \frac{2\rho^2\sigma_\gamma^8}{(1-\rho^2)^4} + \frac{\rho^4\sigma_\gamma^8}{(1-\rho^2)^4}}$$

$$\frac{\left(\frac{\sigma_\gamma^4}{(1-\rho^2)^2} - \frac{\rho^2\sigma_\gamma^4}{(1-\rho^2)^2}\right)}{\frac{\sigma_\gamma^8}{(1-\rho^2)^4} - \frac{2\rho^2\sigma_\gamma^8}{(1-\rho^2)^4} + \frac{\rho^4\sigma_\gamma^8}{(1-\rho^2)^4}}$$

$$-\frac{\left(-\frac{\rho\left(-\beta_0 - \frac{\beta_1 \log[t[1]]}{\log_2} + mm[1]\right)\sigma_\gamma^2}{1-\rho^2} + \frac{\left(-\beta_0 - \frac{\beta_1 \log[t[2]]}{\log_2} + mm[2]\right)\sigma_\gamma^2}{1-\rho^2}\right)\left(\frac{\rho\sigma_\gamma^4}{(1-\rho^2)^2} - \frac{\rho^3\sigma_\gamma^4}{(1-\rho^2)^2}\right)}{\frac{\sigma_\gamma^8}{(1-\rho^2)^4} - \frac{2\rho^2\sigma_\gamma^8}{(1-\rho^2)^4} + \frac{\rho^4\sigma_\gamma^8}{(1-\rho^2)^4}}$$

Consider, now, the problem of representing the probability of the observed values of mm(t) and $\log_2$(age) for a person with missing observations. For example, assume assessments were administered at ages 9, 13, and 17 and the individual is missing the age 13 assessment. The probability density function of the observed errors, $\{\varepsilon[1], \varepsilon[3]\} = \{mm[1] - (\beta_0 + \beta_1 \log_2(t)[1]), mm[3] - (\beta_0 + \beta_1 \log_2(t)[3])\}$, is found by dropping the expected value of the second error element in the PDF statement and the second row and second column in the covariance matrix. In this case, the probability of the observed assessments is given by equation [5],

$$\text{Prob}[\{\varepsilon[1], \varepsilon[3]\}] = ,$$

$$\text{PDF}\left[\text{Multinornal Distribution}\left\{0, 0\right\}, \begin{bmatrix} \frac{\sigma_v^2}{1-\rho^2} & \frac{\rho^2\sigma_v^2}{1-\rho^2} \\ \frac{\rho^2\sigma_v^2}{1-\rho^2} & \frac{\sigma_v^2}{1-\rho^2} \end{bmatrix}\right]$$

$$[\{mm[1] - (\beta_0 + \beta_1 \log_2 t[1]), mm[3] - (\beta_0 + \beta_1 \log_2 t[3])\}]$$

$$\frac{1}{2}\left(-\frac{\left(-\alpha - \frac{\beta \log[t[3]]}{\log_2} + mm[3]\right)(-\beta_0 \log_2 + \beta_0\rho^2\log_2 + \beta_1\rho^2 \log[t[1]]}{(1+\rho^2)\log_2 \sigma_\gamma^2}\right.$$

$$\frac{-\beta_1 \log[t[3]] - \rho^2 \log_2 mm[1] + \log_2 mm[3])}{(1+\rho^2)\log_2 \sigma_\gamma^2}$$

$$-\frac{\left(-\beta_0 - \frac{\beta_1 \log[t[1]]}{\log_2} + mm[1]\right)(-\beta_0 \log_2 + \alpha\rho^2\log_2 - \beta_1 \log[t[1]]}{(1+\rho^2)\log_2 \sigma_\gamma^2}$$

$$\left.\frac{+\beta_1\rho^2 \log[t[3]] + \log[t[3]] + \log_2 mm[1] - \rho^2 \log_2 mm[3])}{(1+\rho^2)\log_2 \sigma_\gamma^2}\right)$$

$$= \frac{1}{2\pi\sqrt{\frac{\sigma_\gamma^4}{(-1+\rho^2)^2} - \frac{\rho^4\sigma_\gamma^4}{(-1+\rho^2)^2}}}$$

Based on these methods, the probability for each person's observed assessments was analytically determined. These probabilities are functions of four parameters, $\{\beta_0, \beta_1, \sigma_v, \rho\}$. The parameters of the problem were estimated with full-information maximum likelihood methods. The log-likelihood of the problem was constructed by taking the logarithm of each probability expression and then these logarithms were added. Gauss-Siedel estimation was used to determine starting values for the parameter estimates (Thisted, 1988). Estimation used *Mathematica*'s Find Maximum command. The estimate of the variance-covariance matrix of the parameter estimates is based on mathStatica's Hessian command (Rose & Smith, 2011; mathStatica, 2011).

» APPENDIX B

This section shows the effect of how small changes in β_1. The first change are a 10% increase in β_1 and then a 20% increase in β_1.

β_1 plus 10% = $(1 + 0.1)\beta_1$

β_1 plus 20% = $(1 + 0.2)\beta_1$

Where,

$\beta_1 = 88.17$

β_1 plus 10% = 96.99

β_1 plus 20% = 105.80

So, in terms of relationship between age and stage with a 10 percent increase in the estimated β_1

Stage = 4.00 + 0.02 MMS

Stage = $4 + 0.02(-73.92 + \beta_1$ plus 10% $*\log_2(\text{age}))$

In terms of relationship between age and stage with a 20 percent increase in the estimated β_1

Stage = $4 + 0.02(-73.92 + \beta_1$ plus 20% $*\log_2(\text{age}))$

The second change is to make a 10% decrease in β_1 and then a 20% decrease in β_1.

β_1 minus 10% = $(1 + 0.1)\beta_1$

β_1 minus 20% = $(1 + 0.2)\beta_1$

Where,

$\beta_1 = 88.17$

β_1 minus 10% = 79.35

β_1 minus 20% = 70.54

So, in terms of relationship between age and stage with a 10 percent decrease in the estimated β_1

Stage = 4.00 + 0.02 MMS

Stage = $4 + 0.02(-73.92 + \beta_1$ minus 10% $*\log_2(\text{age}))$

In terms of relationship between age and stage with a 20 percent increase in the estimated β_1

Stage = $4 + 0.02(-73.92 + \beta_1$ minus 20% $*\log_2(\text{age}))$

» APPENDIX C

This section shows the relationship between change in stage and age. First the relationship between the change in stage and age is shown due an increase in age with a 10 percent increase in the estimated β_1 and then a 20 percent increase in the estimated β_1.

Stage = $4.00 + 0.02 (\beta_0 + \beta_1\log_2(\text{age}))$

$\kappa = D[\text{Stage, age}]$

Where, D is the partial derivative of stage with respective to age, i.e. partial derivative of

$(\beta_0 + \beta_1\log_2(\text{age}))$

$= \dfrac{0.0288539\beta_1}{\text{age}}$

$= 0.02 (\beta_0 + \beta_1 \log_2 (\text{age}))$

$= \dfrac{2.539}{\text{age}}$

So in terms of relationship between age and the change in stage due an increase in age with a 10 percent increase in the estimated β_1

Stage = $4.00 + 0.02 (\beta_0 + \beta_1\log_2(\text{age}))$

$\text{Change in Stage} = 4.00 + 0.02\dfrac{0.0288539\,(\beta_1 + 10\%)}{\text{age}}$

Second, the relationship between age and the change in stage is shown due an increase in age with a 10 percent decrease in the estimated β_1 and then a 20 percent decrease in the estimated β_1.

In terms of relationship between age and the change in stage due an increase in age with a 10 percent increase in the estimated β_1

$\text{Change in Stage} = 4.00 + 0.02\dfrac{0.0288539\,(\beta_1 + 20\%)}{\text{age}}$

In terms of relationship between age and the change in stage due an increase in age with a 10 percent decrease in the estimated β_1

$\text{Change in Stage} = 4.00 + 0.02\dfrac{0.0288539\,(\beta_1 - 10\%)}{\text{age}}$

In terms of relationship between age and the change in stage due an increase in age with a 20 percent decrease in the estimated β_1

$\text{Change in Stage} = 4.00 + 0.02\dfrac{0.0288539\,(\beta_1 - 20\%)}{\text{age}}$

The stage of development of a species predicts the number of neurons

William Joseph Harrigan[1] and Michael Lamport Commons[2]

1 Harvard University
2 Harvard Medical School

ABSTRACT

Does the complexity of an organism's behavior predict the number of neurons in an organism's brain? In the model of hierarchical complexity, the behavioral stage of any organism can be assessed. These behaviors fall into discrete stages. The behavioral stage of development of an organism is defined by the highest order task that an organism has been observed performing. In this study, literature was reviewed to find animals where a neuron count had been taken, and to find behavioral studies to score for stage of development. Once those determinations were made, a power regression analysis addressed the question of whether the behavioral stage of development at which a species operating at predicts the number of neurons an organism has. The relationship between these two variables was $r(17) = 0.874$. These findings imply developing to the next higher stage requires an increase in the number of neurons a species has. The evolutionary benefit from a species evolving to have more neurons may be driven by reinforcement contingencies in the environmental niche that species occupies. If these reinforcement contingencies are one order of hierarchical complexity higher than the stage the species operates at, then the species must increase the number of neuronal connections; this increase reaches a maximum dictated by the number of neurons, so there is a time when the species must evolve more neurons. to perform the comparatively more hierarchically complex tasks required to attain new reinforcement. Therefore it is the attraction of higher stage reinforcers that drives neural development. This neurological correlation for behavioral complexity shows that there is a countable amount of processing power that limits the rate of stage change in a lifetime. The accuracy with which stage of development predicts the number of neurons cast behavioral development as a driving force in neuronal evolution.

KEYWORDS: model of hierarchical complexity, animal behavior, number of neurons, neural development, neuron count, neuron number, behavioral development

A S HAS BEEN SHOWN in a number of studies, brain size differs widely in different species of mammals (see Herculano-Houzel, Collins, Wong & Kaas (2007) for a brief review). Herculano-Houzel and colleagues have also shown that brains of a variety of rodents and primates differ both in terms of their number of neurons and non-neuronal cells (Herculano-Houzel et al., 2007; Herculano-Houzel, Mota & Lent, 2006). These studies showed that the size of primate brains were linearly correlated to the numbers of neurons in those brains, with the size of the cells remaining constant. In rodents, brain size increased faster than the number of neurons. This was due to the finding that in larger rodent brains there was also an increase in neuron size. Their findings show that the number of neurons are better indicators of differences between species than is total surface area or volume of brains. At the same time, the ratio of neuronal cells to nonneuronal cells is constant across studied primates. Although the authors of these papers speculate that these differences in how brains are structured must be related to the increased cognitive abilities of primates as compared to rodents, there has thus far been little actual performance data that is explicitly related to this new measure of the number of neurons in the brains of different organisms.

Note that relating brain measurements such as the number of neurons to how different species perform on a variety of tasks is a difficult one, since many of the traditional measures have been developed and used with only a very limited range of species. For example, intelligence tests, developed on humans, cannot generally be used to assess animals. On the other hand, a great deal of the literature in the learning field uses extremely simple
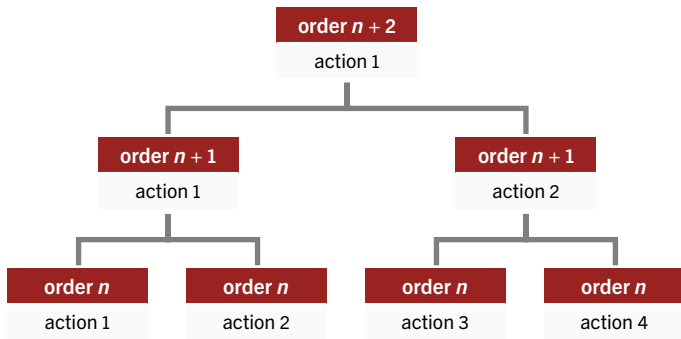
Figure 1. This figure demonstrates the coordination of same-order lower task actions by higher order task actions across two orders of complexity. Starting at the bottom of the figure, four tasks of order *N* are nonarbitrarily coordinated to form two tasks of order *N* + 1 and then two tasks at order *N* + 1 coordinated to form a task at order *N* + 2.

tasks that are appropriate for a variety of animal species but do not particularly tap into the higher levels of human performance. In our research group we have developed and have been working on a measure of performance that can be used across all species of animals, and is not tied to a particular type of task. This is called the *model of hierarchical complexity* (Commons & Pekker, 2008; Commons, et al. 2014).

The model of hierarchical complexity

The model of hierarchical complexity is a model of task complexity. It proposes that tasks can be ordered in terms of their hierarchical complexity using an equally-spaced unidimensional ordinal scale. It is used to predict the difficulty of behavioral tasks independent of domain and content. For a comprehensive review, see Commons, Gane-McCalla, Barker, & Li (2014).

Hierarchical complexity refers to the number of times that the coordinating actions must organize lower order actions. The hierarchical complexity of an action is determined by decomposing the action into the two or more simpler actions that make it up. This iterative process is done until the organization can only be carried out on a set of simple elements that are not built out of other actions. Actions at a higher order of hierarchical complexity can be described by several traits: *1)* They are defined in terms of actions at the next lower order of hierarchical complexity; *2)* organize and transform the lower-order actions; *3)* produce organizations of lower-order actions that are new and not arbitrary, and cannot be accomplished by those lower-order actions alone. Once these conditions have been met, the higher-order action coordinates the actions of the next lower order.

An example of the application of these axioms is shown in Figure 1. Using these axioms, it has been shown that tasks can be categorized into 17 orders of complexity (See Table 1). The order of hierarchical complexity is obtained by counting the number of hierarchical steps, with each step

Table 1. order number and name

order	
number	name
0	computational
1	automatic
2	sensory or motor
3	circular sensory motor
4	sensory-motor
5	nominal
6	sentential
7	preoperational
8	primary
9	concrete
10	abstract
11	formal
12	systematic
13	metasystematic
14	paradigmatic
15	crossparadigmatic
16	meta-crossparadigmatic

Note. Adapted from "Correspondence between some life-span, stage" by Commons, M.L. & Tuladhar, C.T., 2014, *Behavioral Development Bulletin*, *19*(3), p.26.

consisting of a coordination of lower order actions. An organism is said to be operating at a stage when it successfully completes a task at that order of hierarchical complexity (Commons, 2007).

Theory

This paper examines whether the stage of development an organism performs at predicts the number of neurons it has. The hypothesis of this paper is that neural development is an evolutionary means to obtaining higher rates of reinforcement. The mechanism is that successfully performing more hierarchically complex tasks, more reinforcement will be attained. With more reinforcement, the chances of survival should increase. The minimum number of actions required to perform a task at *N* order of hierarchical complexity = 2^N (Commons, Gane-McCalla, Barker, & Li, 2012), so that as task complexity rises as a power function so to should the amount processing power i.e. number of neurons.

It is important to start out by deriving a measure of stage that can be predicted. This starts with the following definition.

Total amount of hierarchical complexity of a task = 2^N_{OHC}.

This is based on Pascual-Leone's (1970) suggestion that to solve a problem at order *N*, there needs to be a working memory of 2^N. Due to the definitions of stage and of order given in Step 1, the stage number, N_{stage}, is the same as the number N_{OHC}, for the most hierarchically complex task solved

$$2^N_{OHC} = 2^N_{stage}.$$

Therefore, what we will be predicting is N_{stage} or performance. It is also predicted that because stage is 2 to a power, the function relating stage to the number of neurons will be a power function.

The above table shows all known orders of hierarchical complexity and their corresponding numbers.

» METHOD

This paper investigates the relationship between the number of neurons and behavioral developmental stage attained in adults of a species. The number of neurons an animal had was obtained from published work. To assess stage, the highest order of hierarchical complexity of tasks completed by an adult of the species was assessed (*N* = 19). The order of hierarchical complexity of the task that each species has successfully completed was obtained from published work describing whichever task has been characteristically used with that species. This was done for eighteen different species for which adequate data on the tasks completed by that species or a highly similar species was available. The stage of an animal's performance on a task was defined as successful performance on that task of the same order. The stage data was found by scoring the tasks. The average numbers of neurons that has been measured in each species was found in published literature.

This paper predicts that the relationship between stage and number of neurons will be better expressed as a power function than as a linear function. This was tested using both linear regressions and non-linear regression, with stage of development as the independent variable and number of neurons as the dependent variable.

Animals species

Below is a list of species for which the investigator ascertained the neuron counts and had task descriptions. For each species or species group there had to be a published account of its behavior, which the investigators scored for stage of development. As can be seen, each species is placed within the stage that their behavior was scored at.

The weights of different species ascertained. When the sources for the weights in question yielded a range of weights, the mean of that range was used. When different weights for male and female members of the same species were found, the mean of those two number was used.

Automatic stage 1

At the automatic order 1, a single action that is an innate biological response to a single environmental stimulus. This stimulus is not paired with any other stimulus. Examples of the environmental stimulus could be a chemical emitted by possible food, or a physical stimulus such as light. The actions are "hard wired" into the organism. Examples include taxis, tropisms, phagocytosis and unconditionable reflexes. The organisms that perform these actions are single celled. While habituation and sensitization occur at this stage, they are not coordinated into classical conditioning (Commons & Giri, 2014). No organisms with brains have been found that operate with automatic stage 1 as their highest stage, therefore no stage 1 organisms were included in these analyses.

Sensory or motor stage 2

Respondent conditioning at order 2 of hierarchical complexity coordinates two stimulus response pairs from the lower automatic order 1. Two characteristics of this order are: *a)* two stimuli are paired either in a naturalistic environment or by an experimenter. In other words, an unconditioned stimulus that already elicits an unconditioned response is paired with another salient stimulus and, *b)* the organism's behavior does not directly cause the reinforcing stimuli in this situation as it does in operant conditioning. Reflexes that are conditioned are also order 2 behaviors.

Caenorhabditis elegans			
stage:	2	average weight:	(not found)
neurons:	302[†]	neurons per gram:	(not available)

[†]*White et. al., 1986*

Caenorhabditis elegans, a kind of roundworm, has been classically conditioned in a laboratory environment (Rankin, 2000). No papers were found stating that this species has performed any tasks more hierarchically complex than being classically conditioned. Therefore the highest stage observed is stage 2, and this organisms' neuron count will be used.

Circular sensory-motor stage 3

Operant conditioning is an order 3 action. Operant conditioning is built out of the non-arbitrary coordination of three sensory or motor order 2 task actions or steps. These steps are step 1, "What to do"; step 2, "When to do it"; and step 3, "Why to do it" (Commons & Giri, in press). The three steps of respondent conditioning are from order 2 but are not coordinated until order 3.

Three very different cases of procedural respondent conditioning are used. The only commonality between the three respondent conditioning steps is the basic procedure. Those procedural steps are the "What to Do" (step 1), "When to Do" (step 2) and "Why to Do" (step 3). In step 1 of the respondent conditioning the *representation of behavior* takes on the elective properties of the s^{R+} making the *representation of behavior* salient. We leave the representation of behavior undefined. One might use common notions of it instead. In step 2, the now salient *representation of behavior* (*rb*) is paired with an environmental *s*. This makes the *s* elicit the representation of a behavior which requires the saliency of the *representation of a behavior*. In step 3, the environmental *s* is paired with the s^{R+} making the *s* more salient and valuable. When the environmental stimulus is more salient, the representation of a behavior rate relative to other *representation of a behavior*'s not associated with reinforcement increases.

Drosophila melanogaster (fruit flies)			
stage:	3	average weight:	~0.25 milligrams[‡]
neurons:	~100,000[†]	neurons per gram:	400,000,000

[†]*Bharucha, 2009;* [‡]*Seiger, 1965*

Operant conditioning is the product of non-arbitrarily coordination of two or more classical conditioning contingencies (Commons & Giri, submitted). *Drosophila melanogaster* have been operantly conditioned in a laboratory environment (Brembs & Heisenberg, 2000), so they have performed an order 3 task. No order 4 tasks were found in a literature search. Therefore the highest stage observed is stage is 3.

Aplysia californica			
stage:	3	average weight:	(not found)
neurons:	20,000[†]	neurons per gram:	(not available)

[†]*Retrieved November 9, 2014 from the Broad Institute website.*

Aplysia Californica have been operantly conditioned by pairing electric stimulating the anterior branch of the esophageal nerve (En2) with biting behavior (Baxter & Byrne 2006).

Apis mellifera (honey bee)			
stage:	3	average weight:	0.1 grams[‡]
neurons:	960,000[†]	neurons per gram:	9,600,000

[†]*Menzel & Giurfa, 2001;* [‡]*Needham, 1995*

Erber, Pribbenow, Kisch, & Faensen (2000) trained honeybees (*Apis mellifera*) to move their antenna muscles to receive sugar water. This coordinates an arbitrary muscle motion with a reward. This simultaneously associatively conditions the bee to associate the touching of the stimuli objects with the reward of sugar water, in that non-arbitrary order. This pairs two order 2 behaviors into an order 3 task.

The only challenge to stage 3 being the ceiling of the stage that honeybees reach is the waggle dance (Von Frisch 2011). The waggle dance is a communicative act that begins when one bee returns to the hive after finding some food. This food source may be more than a kilometer away. This bee moves into the vicinity of other bees and moves its body in very particular ways. After this waggle dance is completed the bees who witnessed it navigate their way to the aforementioned food with few errors.

This behavior of following the directions from the waggle dance is a non-arbitrary sequence of operant (stage 3) behaviors, however these behaviors consist of simply following a sequence of motions over time. This paper argues that giving or following such a sequence is long chain of order 3 tasks, rather than an order 4 task. This argument is supported by the notion that the waggle dance may not be learned behavior. Ai and Hagio, (2013) found evidence that there is considerable specialization in bee anatomy that assists in the performance and response to this dance. In this case it would be an operant procedure requiring a tremendous amount of sequential coding. For the purposes of this paper, the waggle dance is a very horizontally complex order 3 task.

Rana esculenta (frog)			
stage:	3	average weight:	50 grams[‡]
neurons:	16,000,000[†]	neurons per gram:	320,000

[†]Holmes, 1969; [‡]retrieved December 31, 2014, from http://a-z-animals.com/animals/common-frog/

Neveu, (2011) found that *Rana esculenta* could be trained to eat pellets, despite the fact that in the wild they only eat things that move. In this process the frogs are performing not only the order 2 task of learning to treat new substances as food, but also a whole new order 3 eating procedure.

Sensory-motor stage 4

At sensory-motor, order 4, organisms coordinate 2 or more circular sensory-motor subtask actions into a superordinate "concept." New and untrained instances of the concept are responded to correctly. These correct responses do not depend on simple stimulus generalization.

Rat			
stage:	4	average weight:	337 grams[‡]
neurons:	200,000,000[†]	neurons per gram:	592,592

[†]Herculano-Houzel, 2005; [‡]retrieved December 31, 2014, from http://web.jhu.edu/animalcare/procedures/rat.html

The following is a description of order 4 behavior in rats. Rats were repeatedly presented with three scented stimuli. Two were always of identical scent, while the third was always different from the other two. The scents were different every trial. Reinforcement was received for selecting the third stimulus that was scented differently from the other two (Bailey & Thomas 1998). They had to discriminate what is termed oddity matching. This is an order 4 task, which coordinates multiple order 3 operant contingent behaviors. A literature search did not find any more hierarchically complex tasks than this performed by rats, therefore rat are operating at stage 4.

Mouse			
stage:	4	average weight:	26.5 grams[‡]
neurons:	71,000,000[†]	neurons per gram:	2,679,245

[†]Herculano-Houzel, Mota, & Lent, 2006; [‡]retrieved December 31, 2014, from http://web.jhu.edu/animalcare/procedures/mouse.html

Watanabe (2013) demonstrated that mice can be conditioned to discriminate between the paintings of different artists. This process coordinates multiple order 3 (operant) cues to see the conceptual differences between paintings.

The example behaviors for rats were originally obtained from an unpublished paper by Miller, Commons, Commons-Miller, and Chen (2014).

The behavior of the following rodents was deemed to be similar enough to other rodents that the same is used for them:

Agouti			
stage:	4	average weight:	4200 grams[‡]
neurons:	857,000,000[†]	neurons per gram:	204,048

[†]Herculano-Houzel, Mota, & Lent, 2006; [‡]Emmons, 1997

Capybara			
stage:	4	average weight:	50,500 grams[‡]
neurons:	1,600,000,000[†]	neurons per gram:	31,683

[†]Herculano-Houzel, Mota, & Lent, 2006; [‡]Macdonald, 2006

Hamster			
stage:	4	average weight:	200 grams[‡]
neurons:	89,970,000[†]	neurons per gram:	449,850

[†]Herculano-Houzel, Mota, & Lent, 2006; [‡]retrieved December 31, 2014, from http://hamsters-uk.org/content/view/70

Guinea pig			
stage:	4	average weight:	950 grams
neurons:	89,970,000[†]	neurons per gram:	252,232

[†]Herculano-Houzel, Mota, & Lent, 2006

Tree shrew			
stage:	6	average weight:	190 grams[‡]
neurons:	261,400,000[†]	neurons per gram:	1,375,789

[†]Herculano-Houzel, Collins, Wong, & Kaas, 2007; [‡]Payne Francis, Phillips, 1985

Nominal stage 5

Characteristics of order 5 include responding to words that represent concepts. They also follow sequences of word commands. A single word command is stage 4, sequences of them is stage 5.

A literature search was performed for dogs and cats, both are animals hypothesized to be performing at stage 5. No reliable data for the number of neurons in the whole brain was found.

Sentential stage 6

A characteristic of order 6 is following sequences of stage 5 representations of concepts. An example of this in humans is constructing sentences. This is the earliest form of grammar.

A literature search was performed for Crows, and African Grey parrots, both are animals previously informally scored to be performing at stage 5. No reliable data for the number of neurons in the whole brain was found for these animals. Some kinds of monkeys, however, can be scored at this stage.

Except when noted otherwise, weights for the primates in this section were retrieved from the University of Wisconsin Madison Primate Info Net.

Capuchin monkeys			
stage:	6	average weight:	3085 grams
neurons:	3,690,000,000[†]	neurons per gram:	1,196,110

[†]Herculano-Houzel, Collins, Wong, & Kaas, 2007

Chen, Lakshminarayanan, and Santos (2005) performed a study where Capuchin monkeys were given tokens to trade for rewards. In this study, the Capuchin participants successfully exchanged tokens for different food rewards based of the preferences of the individual participant, and changed exchange rate of token for the different kinds of food rewards. The fact that the Capuchin monkeys changed their buying habits in response to changes in price shows that they can accurately respond multiple values for the same token. Using a token to represent a single concept is an order 6 task. Using the same token differently in different contexts coordinates multiple order 5 tasks into an order 6 task.

The behavior of the following small non-ape primates was deemed to be similar enough to squirrel monkeys that the same scoring could apply to these animals:

Squirrel monkey			
stage:	6	average weight:	861.5 grams
neurons:	3,246,430,000[†]	neurons per gram:	3,768,346

[†]Herculano-Houzel, Collins, Wong, & Kaas, 2007

Marmoset			
stage:	6	average weight:	246 grams
neurons:	635,800,000[†]	neurons per gram:	2,584,553

[†]Herculano-Houzel, Collins, Wong, & Kaas, 2007

Galago			
stage:	6	average weight:	1250 grams[‡]
neurons:	936,000,000[†]	neurons per gram:	748,800

[†]Herculano-Houzel, Collins, Wong, & Kaas, 2007
[‡]O'Mara, Gordon, Catlett, Terranova, & Schwartz, 2012

Owl monkey			
stage:	6	average weight:	937.25 grams
neurons:	1,468,000,000[†]	neurons per gram:	1,566,284

[†]Herculano-Houzel, Collins, Wong, & Kaas, 2007

Preoperational stage 7

Organisms form lists of organized sets of acts and make simple deductions that connect simple sequences of actions (without contradiction excluded). A human telling a story, for example, is like a sequence of sentences. One of the end results includes that organisms can count random events and objects placed in a row or presented in a sequence, combine numbers, and combine simple propositions.

Rhesus monkey			
stage:	7	average weight:	5300/7700 grams[‡]
neurons:	6,380,000,000[†]	neurons per gram:	981,538

[†]Herculano-Houzel, 2011; [‡]male/female

Washburn and Rumbaugh (1991) trained Rhesus monkeys to select Arabic numerals associated with a number of food pellets. This task coordinates the order 6 sequence of numerals with the order 6 sequence of numbers of objects to create an order 7 action. A literature search yielded no order 8 tasks, so they are scored as operating at order 7.

Primary stage 8 and concrete stage 9

Logical deduction and empirical rules are applied in the primary order. In concrete order 8, simple logical deduction and time sequences are used to describe actual instances. The instances are actual because they occur in past or present time. They are composed of specific things, incidents events, actions, actors and places. Concrete order 8 actions are applied to a small number of specific instances.

Chimpanzees

stage:	8	**average weight:**	44,000 grams
neurons:	26,040,000,000[†]	**neurons per gram:**	581,899

[†] *estimated number of neurons based on brain weight and a model of primate data (Herculano-Houzel, 2009)*

Chimpanzees were not included in the analysis because the investigators did not find an empirical neuron count for them. The number of neurons given above is just an estimate.

Gomes and Boesch found that chimpanzees engage in a variety of trading behaviors including the exchange of meat, social support, and sex. The appraisal of value for a good or service is an order 7 task. To make a deal requires the non-arbitrary coordination of two or more such values, and it is therefore an order 8 task. No published evidence of order 9 behavior in chimpanzees has been found, so chimpanzees are scored to be operating at stage 8.

Table 2. every animal in this analysis organized by number of neurons

species	stage	neurons	grams	neurons/grams
human	11	86,060,000,000	62,000	1,388,065
rhesus monkey	7	6,380,000,000	6,500	981,539
capuchin monkey	6	3,690,000,000	3,085	1,196,110
squirrel monkey	6	3,246,430,000	861.5	3,768,346
capybara	4	1,600,000,000	50,500	31,683
owl monkeys	6	1,468,000,000	937.25	1,566,284
galago	6	936,000,000	1,250	748,800
agouti	4	857,000,000	4,200	204,048
marmoset	6	635,800,000	246	2,584,553
tree shrew	6	261,400,000	190	1,375,789
guinea pig	4	239,620,000	950	252,232
rat	4	200,000,000	337.5	592,592
hamster	4	89,970,000	200	449,850
mouse	4	71,000,000	26.5	2,679,245
frog	3	16,000,000	50	320,000
honey bee	3	960,000	.1	9,600,000
fruitfly	3	100,000	0.00025	400,000,000
Aplysia californica	3	20,000	not found	—
C. elegans	2	302	not found	—
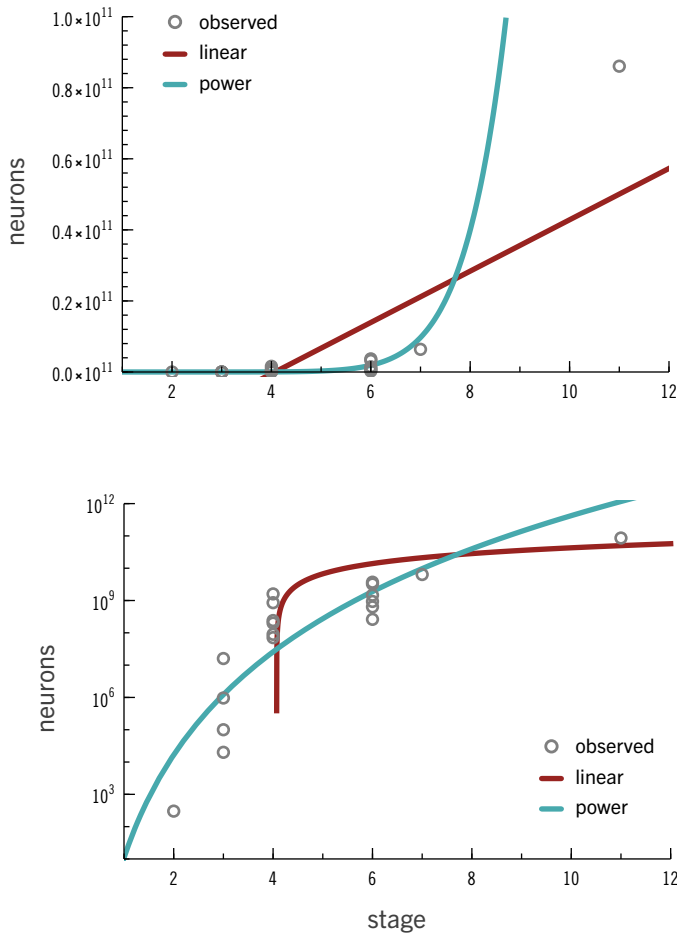
Figure 2. Both of these plots show the results of the regression of number of neurons on stage of development, though the bottom one is using log-transformed coordinates. Each circle represents a species. The *y*-axis shows the number of neurons and the *x*-axis shows the developmental stage of the species. The solid line shows a linear function, and the dotted line shows a power function.

Abstract stage 10, formal stage 11, systematic stage 12

Humans

stage:	10—12	**average weight:**	62,000 grams[†]
neurons:	86,060,000,000	**neurons per gram:**	1,388,065

[†] *Walpole, Prieto-Merino, Edwards, Cleland, Stevens, & Roberts, 2012*

Human beings have a much wider range of developmental stages than any other animal. In Commons et al. (2005), at stage 10, humans coordinate a number of concrete instances of events, example, etc. to form variables. This allows for relative values that include such as stereotypes, ingroups and outgroups. Also emergent at this stage are variable quantities and qualities, as well as categorical assertions.

At formal stage 11, humans coordinate two variables into one-dimensional linear logic. Analytic examples of this include syllogistic logic, and univariate algebra. In univariate algebra, simple equations with one unknown (a variable) are solved. One gets a relationship between *y*, the dependent variable and *x* the independent variable.

At systematic stage 12 humans coordinate multiple stage 11 relationships among variables tasks. An example of this is solving systems of equations. (Commons et. al., 2005)

Not only are humans born at stage one, but stages 9-15 have only been observed in humans. Very few (less than 2%) of humans have been found to perform at stage 13 or above. Because of this special situation, the mean stage of stage 11 was used in the analysis (Commons, & Ross, 2008; Commons, Li, Richardson, Gane-McCalla, Barker, & Tuladhar, 2013).

The above table shows every animal in this analysis organized by number of neurons. From right to left is the name of the species, the highest stage the species was scored as operating at, the number of neurons the species has, the average weight of the species in grams, and the number of neurons the species has divided by their weight in grams.

» RESULTS

Regression of number of neurons and the stage of species

There were two groups of regressions performed. Each group used a different dependent variable: In the first group, the dependent variable was number of neurons; and in the second group the dependent variable was number of neurons per gram of body weight. For each group, first a traditional linear regression was used. And then second, a power function regression was used. This was because it was expect that a Power Function would fit much better because the number of actions required for a given stage was 2^N where N is the stage number.

Equations underlying regression

The following shows the equations underlying the regression analyses performed in this study. First, the linear regression equation is

$$y = \beta_0 + \beta_1 x, \tag{1}$$

where y is the number of neurons, x is the stage of performance of the animal, β_0 is the y-axis intercept constant in the linear case and the multiplicative constant in the power case, and β_1 is the slope constant in the linear case and the exponent in the power case.

With all values plugged in, the linear equation looks like:

$$y = 2.942 \times 10^{10} + 7.266 \times 10^9 x. \tag{2}$$

Second, the power regression equation is

$$y = \beta_0 x^{\beta_1}. \tag{3}$$

With all the values plugged in, the power equation looks like this

$$y = \beta_0 x^{\beta_1} = 7.266 \times 10^9 x^{-2.942 \times 10^{10}}. \tag{4}$$

Results of regression analyses

First, a linear regression was performed to examine the question of whether the average total number of neurons in a species predicts the highest stage of development observed in that species (See Figure 2). The results were $r(17) = 0.762$ ($R^2 = 0.580$, $p < 0.001$). This shows a strong relationship between number of neurons and the complexity of task observed in the animal.

As predicted, a power function predicted the number of neurons better, $r(17) = 0.874$ ($R^2 = 0.764$, $p < 0.001$) (See Figure 2). The animals with extremely high and low numbers of neurons are quite different from those in the middle
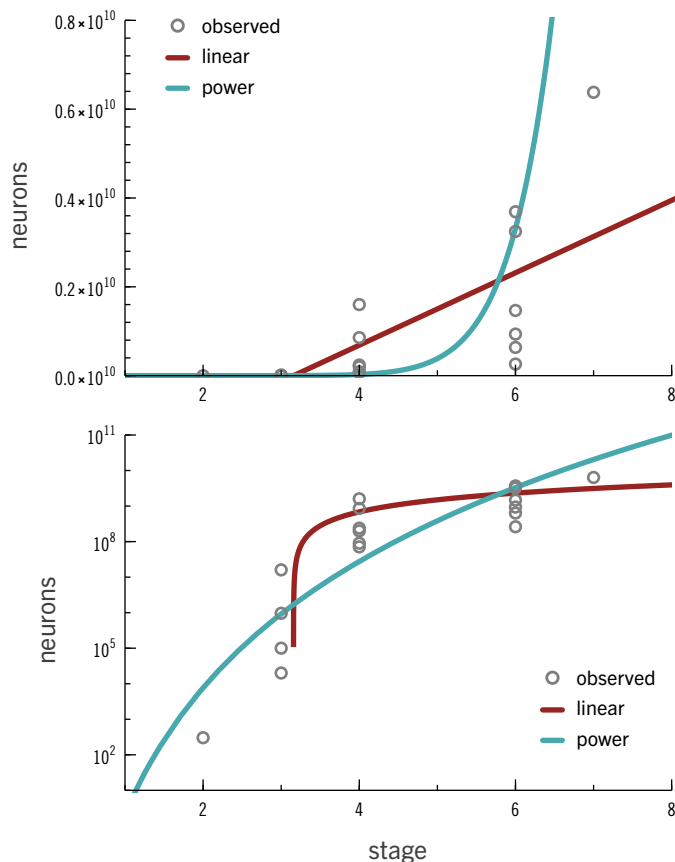
Figure 3. Both of these plots show the results of the regression of number of neurons on stage of development (not including humans), though the bottom one is using log-transformed coordinates. Each circle represents a species. The y-axis shows the number of neurons and the x-axis shows the developmental stage of the species. The solid line shows a linear function, and the dotted line shows a power function.

Regression of number of neurons on the stage of the species without humans

In order to investigate whether the extremely high stage of humans was distorting the regression, the investigators performed an additional regression that did not include humans. See Figure 2. The linear regression without humans resulted in a slightly higher $r(16) = 0.695$ and $R^2 (16) = 0.483$ ($p < 0.001$). The power regression without humans resulted in a slightly higher $r(16) = 0.871$ and $R^2 (16) = 0.759$ ($p < 0.001$). In summary removing humans from the regression analysis did not significantly change the results of the analysis. This figure demonstrates that there is a strong relationship between stage and neurons even when humans are not included in the analysis.

Regression of number of neurons divided by weight from the stage of the species

The next series of analyses examined the question of whether the stage of development of the species this paper discusses can be predicted by the number of neurons per gram of body weight. To test this hypothesis, stage of development of the organism divided by the weight of the organism was used to predict number of neurons. *Aplysia californica* and *Caenorhabditis elegans* were not included in this analysis, because their weight was not included. (See Figure 3).

The resulting $r(15) = 0.276$ ($R^2 = 0.077$, $p = 0.142$) was quite low. The power regression was worse $r(15) = 0.107$ ($R^2 = 0.012$, $p = 0.341$).

Stage did not predict the number of neurons to body weight ratio as well as it predicted the simple number of neurons. This may be seen by comparing the results to the other regression fits in this paper. These preliminary results support the hypothesis outlined by Herculano-Houzel (2011) that the total amount of neurons is more predictive of performance than brain size, or brain to body ratio.

Using stage of development to predict the number of neurons

The following shows how the equations underlying regression can be used to predict the number of neurons for an organism that is operating at a particular stage. As stated earlier in the results section the equation for linear regression is as follows. As stated earlier in this section, the linear prediction of how many neurons one expects to find for a stage of an animal is shown by

$$y = 2.942 \times 10^{10} + 7.266 \times 10^9 x. \tag{5}$$

Next, it will be shown how to obtain the number of neurons from stage. As stated earlier in this section, the power equations is

$$y = \beta_0 x^{\beta_1}. \tag{6}$$

Within *ln* (natural log) form, this is

$$
\begin{aligned}
\ln(y) &= \ln(\beta_0) + \beta_1 x \\
y &= -2.942 \times 10^{10} + 7.266 \times 10^9 x,
\end{aligned} \tag{7}
$$

where $\ln(y)$ is the natural log of the number of neurons and $\ln(\text{stage})$ and $\ln(\text{neurons})$ are the variables.

$$
\begin{aligned}
\beta_1 &= -2.942 \times 10^{10} \\
\beta_0 &= 7.266 \times 10^9
\end{aligned}
$$

With the estimates put into the equation z

$$\ln(y) = 2.153 + 11.806x. \tag{8}$$

As stated earlier in this section, put this in power function form this then becomes

$$y = \beta_0 x^{\beta_1} = 7.266 \times 10^9 x^{-2.942 \times 10^{10}}. \tag{9}$$

Table 3. the values for the estimated parameters of equation (6) being used to predict the number of neurons an organism has

stage	β_1 slope $= -2.942 \times 10^{10}$	β_0 constant $= 7.266 \times 10^9 x$	predicted neurons
1	−29,420,000,000	7,266,000,000	−22,154,000,000
7	−29,420,000,000	50,862,000,000	21,442,000,000
11	−29,420,000,000	79,926,000,000	50,506,000,000

Note. There cannot be a negative number of neurons which is what the constant β_0, implies.
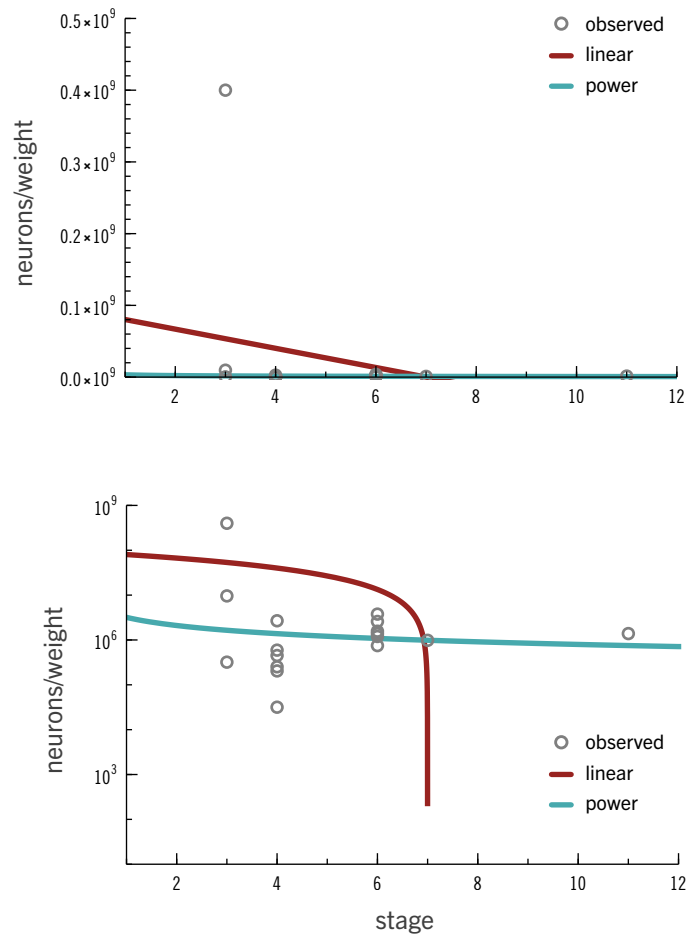
Figure 4. Both of these plots show the results of the regression of number of neurons divided by average weight on stage of development, though the bottom one is using log-transformed coordinates. Each circle represents a species. The *y*-axis shows the number of neurons and the *x*-axis shows the developmental stage of the species. The solid line shows a linear function, and the dotted line shows a power function.

Standard errors

The standard error of the parameters for the analysis that included all the species is given. Note that the standard error of the estimate of β_0 and β_1 are small compared to the size of the estimate and therefore the *p* values are significant.

The Standard Errors of $\beta_1(x) = 1.54 \times 10^9$, $p < 0.001$

The Standard Errors of $\beta_0 = 8.191 \times 10^9$, $p = 0.002$

Finally the standard error of the parameters in natural log form is given. Not that the standard error of the estimate of β_0 is very large and therefore the *p* value shows it is insignificant

The Standard Errors of $\beta_1 \ln(x) = 1.465$, $p < 0.001$

The Standard Errors of $\beta_0 = 21.274$, $p = 0.667$

» DISCUSSION

The most basic finding in this paper is that higher stage behavior requires more neurons. This means number of neurons in the brain limits development of an organism, because the number of neurons limits the number of synapses. The model of hierarchical complexity and the stage of the organism predict the number of neurons rather well.

No attempt is made to explain why there is variability in the number of neurons. A result of stage change is that it increases the chances of survival of an organism. This is because a higher stage organism can address more tasks. As the organisms on earth evolve to greater sophistication every aspect of every organism is subject to mutation. Sometimes brains mutate to have a greater number of neurons. If the organism or its offspring find a way to use these newer neurons to their evolutionary advantage, then the larger number of neurons succeeds and may eventually become the norm. Whereas if these neurons are never put to better use than a smaller amount of neurons then eventually the wasted cells will be selected against and the brain mutation will be lost. Therefore it is not enough to simply have more neurons to raise stage, but rather the organism that have found ways use incremental increases in neurons to find more advantageous behaviors throughout their evolutionary history.

Stage and energy rate density

Another central implication of this research relies on the finding that for most animals, neurons use roughly the same amount of energy per neuron. The amount of energy consumed per neuron across 6 species of rodent and primates varying in size from mice to humans varied by 40% (Herculano-Houzel, 2011). This shows the energetic consumption of the processing power of living organisms may be approximated by counting the number neurons. For a study of this scale data for more species may be required before calculations of energy rate density can be accurately performed. There has been considerable work in energetic complexity as energy rate density (Chaisson, 2012). This paper implies that the neuron could be used as a unit of hierarchical energetic complexity. That means that the stage of problem solving requires more energy as the stage increases. What is found is that the higher the stage, the more effective the animal is in obtaining total amount of energy. This rise in rate of reinforcement in turn increases energy consumption of the organism, and by this mechanism the organisms evolves to be more energetically complex, in term of control over energy in an organism's environment, and neuronal energy consumption.

Number of neurons and working memory

Next an explanation is given why stage change drives the number of neurons. Comparisons can be made between developmental stage and working memory. Developmental stage appears to be contingent on working memory. Working memory in this case is being used in the Pascual- Leone (1970) sense. In a simplified view of how hierarchical complexity and therefore stage interacts with number of neurons, we start with how order of hierarchical complexity is calculated. Higher order task actions are made from the non-arbitrary coordination of 2 or more lower order task actions. Therefore, the total number actions that need to be coordinated is 2^N, where N is the order of hierarchical complexity. Note that this is a count of actions irrespective of their order. There has to be enough "working memory" to correctly address a task, though this working memory is not directly comparable to computer memory. In order for an organism to coordinate lower order tasks into a task at the next higher stage the organism has to store and use the solutions from the lower stages long enough for the coordination to occur. Because of randomness and mutations in the variation in number of neurons there has to be enough neurons to make it possible to solve the next order problem. Working memory also includes the other requirement of the hierarchical organization. The higher order neural networks have to read the lower order neural networks.

Note that the analysis in this paper does not attend to the differences in brain structure between species. It looks only at the number of neurons. However, the results are strongly in support of the hypothesis that stage of development predicts the number of neurons in organisms. It was found that stage predicted number of neurons better than it predicted the neuron-weight ratio.

A power function proved to be more predictive/strongly correlated $r(17) = 0.860$ ($R^2 = 0.740$) than a linear model $r(16) = 0.839$ ($R^2 = 0.704$). The amount of neurons required to perform a task increases as a power function. This follows, as the number of task actions is $= 2^N$, which itself is a power function.

A more robust result could be attained by having more animals. Unfortunately, neuron counts have only begun just nine years ago (Herculano-Houzel, S. & Lent, R. 2005). Ideally, there would be at least two species for every stage and different taxonomic classes for each set of two (C. D. Barker, personal communication, December 30, 2014). For example, this analysis would benefit from having bird neuron counts, and the addition of species can further refine the model.

» CONCLUSION

Stage was used to predict number of neurons. A relatively large $r = 0.874$ was obtained. This high correlation may be interpreted as unidirectional with increase in stage leading to an increase in the number of neurons, and not the other way around. This is because higher order of hierarchical complexity tasks must exist in the environment in order for animals to obtain reinforcement for doing them. It is hypothesized that higher stage behavior afforded by neural development is an evolutionary *means* to obtaining higher rate of reinforcement. The mechanism is that successfully performing more hierarchically complex tasks, more reinforcement will be attained. With more reinforcement, the chances of survival should increase. ■

REFERENCES

Ai, H., & Hagio, H. (2013). Morphological analysis of the primary center receiving spatial information transferred by the waggle dance of honeybees. Journal of Comparative Neurology, 521 (11), 2570-2584.

Azevedo, F.A., Carvalho, L.R., Grinberg, L.T., Farfel, J.M., Ferretti, R.E., Leite, R.E., Filho, W.J., Lent R. & Herculano-Houzel, S. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. Journal of Comparative Neurology, 513 (5), 532-541.

Bailey A.M. & Thomas R.K. (1998). An Investigation of Oddity Concept Learning by Rats. The Psychological Record, 48,333-344

Baxter, D.A., & Byrne, J.H. (2006). Feeding behavior of Aplysia: a model system for comparing cellular mechanisms of classical and operant conditioning. Learning & Memory, 13 (6), 669-680.

Bharucha K.N. (2009) The Epicurean Fly: Using Drosophila Melanogaster to Study Metabolism. Pediatric Research 65 (2), 132-137

Boinski, S. (1991). The coordination of spatial position: a field study of the vocal behaviour of adult female squirrel monkeys. Animal Behaviour, 41 (1), 89-102.

Brembs B. & Heisenberg M. (2000). The Operant and the Classical in Conditioned Orientation of Drosophila Melanogaster at the Flight Simulator. Learning and Memory, 7 (2), 104-115.

Broad Institute. (n. d.). Retrieved November 9, 2014 from the Broad Institute website: http://www. broadinstitute. org/science/projects/mammals-models/vertebrates-invertebrates/aplysia/aplysia-genome-sequencing-project.

Brzoska, J. (1982). Vocal response of male European water frogs (Rana esculenta complex) to mating and territorial calls. Behavioural Processes, 7 (1), 37-47.

Burton, M., & Burton, R. (1974). The Funk & Wagnalls Wildlife Encyclopedia: Vol 1 AAR-BAD, Vol 9 HAR-HUT. Funk and Wagnalls, Incorporated

Chaisson, E.J. (2012). A singular universe of many singularities: Cultural evolution in a cosmic context. In Singularity Hypotheses (pp. 413-440). Springer Berlin Heidelberg.

Chen, M.K., Lakshminarayanan, V., & Santos, L. (2005), The Evolution of Our Preferences: Evidence from Capuchin Monkey Trading Behavior. Cowles Foundation Discussion Paper No. 1524. Available at SSRN: http://ssrn. com.ezp-prod1.hul.harvard.edu/abstract=675503

Common Frog. (n.d.). Retrieved December 31, 2014, from http://a-z-animals. com/animals/common-frog/

Commons, M.L., Gane-McCalla, R., Barker, C.D., & Li, E.Y. (2012). The Model of Hierarchical Complexity as a measurement system. Journal of Adult Development. Retrieved from: http://adultdevelopment.org/CommonsEtAl2012.pdf

Commons, M.L., Li, E.L., Richardson, A.M., Gane-McCalla, R., Barker, C.D., & Tuladhar, C.T. (2013). Does the Model of Hierarchical Complexity Produce Significant Gaps between Orders and Are the Orders Equally Spaced?. Journal of applied measurement, 15 (4), 422-449.

Commons, M.L., Miller, P.M., Goodheart, E.A., & Danaher-Gilpin, D. (2005). Hierarchical Complexity Scoring System (HCSS): How to Score Anything. Unpublished Scoring Manual Available from Dare Institute, Commons@tiac.net http://dareassociation. org/Papers/Scoring%20Manual. htm

Commons, M.L., & Ross, S.N. (2008). What postformal thought is, and why it matters. World Futures, 64(5-7), 321-329.

Commons, M.L. (2007) Introduction to the Model of Hierarchical Complexity Behavioral Development Bulletin 13, 1-4

Commons M.L. & Giri S.P. (2014) Account of Operant Conditioning Based on Coordinating Three Procedural Steps of Respondent Conditioning Processes. Manuscript submitted for publication.

Emmons, L.H. (1997). Neotropical Rainforest Mammals. Pp. 226-230. 2nd edition. Chicago, IL: University of Chicago Press

Erber, J., Pribbenow, B., Kisch, J., & Faensen, D. (2000). Operant conditioning of antennal muscle activity in the honey bee (Apis mellifera L.). Journal of Comparative Physiology A, 186 (6), 557-565.

Hamster Health Checks. (2008). Retrieved December 31, 2014, from http://hamsters-uk.org/content/view/70

Herculano-Houzel, S. & Lent, R. (2005). Isotropic fractionator: a simple, rapid method for the quantification of total cell and neuron numbers in the brain. J Neurosci 25 (10), 2518–2521. doi:10. 1523/jneurosci. 4526-04. 2005

Herculano-Houzel (2011). Scaling of Brain Metabolism with a Fixed Energy budget per Neuron: Implications for Neuronal Activity, Plasticity and Evolution, PLOS ONE 6 (3), e17514

Herculano-Houzel, S., Mota, B., & Lent, R. (2006). Cellular scaling rules for rodent brains. Proceedings of the National Academy of Sciences, 103 (32), 12138-12143.

Herculano-Houzel, S., Collins, C.E., Wong, P., & Kaas, J.H. (2007). Cellular scaling rules for primate brains. Proceedings of the National Academy of Sciences, 104 (9), 3562-3567.

Holmes, R.L. (1969). Atlas of the Frog's Brain. Journal of anatomy, 105(Pt 3), 586.

Macdonald, D.W. (2006) The Encyclopedia of Mammals. Oxford, UK: Oxford University Press.

Menzel, R., & Giurfa, M. (2001). Cognitive architecture of a mini-brain: the honeybee. Trends in cognitive sciences, 5 (2), 62-71.

Miller P.M., Commons M.L., Commons-Miller L, & Chen S. (2014) A Comparative Account of Numerosity in Human and Non-Human Animals Based on the Model of Hierarchical Complexity Manuscript in preparation.

Needham, A. (1995). Frequently Asked Questions About Honey Bees. Retrieved December 31, 2014, from http://www.bees-online.com/BeeFaqs.htm

Neveu, A. (2011). Influence of genotype of froglets belonging to the Rana esculenta hybridogenetic complex in relation to learning capacity to eat pellets. Aquaculture, 310 (3), 343-349

Nieuwenhuys, R., Ten Donkelaar, H.J. and Nicholson (1998), C. The Central Nervous System of Vertebrates. Vol. 3, Berlin: Springer.

O'Mara, M.T., Gordon, A.D., Catlett, K.K., Terranova, C.J., & Schwartz, G.T. (2012). Growth and the development of sexual size dimorphism in lorises and galagos. American journal of physical anthropology, 147 (1), 11-20.

Payne J., Francis, C.M., Phillips, K. (1985) A Field Guide to the Mammals of Borneo, Malaysia. pp. 161–162. Sabah, Malaysia: The Sabah Society.

Pascual-Leone J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. Acta Psychol. 63, 301–345.

Primate Factsheets. (n.d.). Retrieved December 31, 2014, from http://pin. primate.wisc.edu/factsheets

Rankin C.H. (2000). Context Conditioning in Habituation in the Nematode Caenorhabditis elegans. Behavioral Neuroscience, 114 (3), 496-505

Seiger M.B. (1966) The Effects of Chromosome Substitution on Male Body Weight of Drosophila Melanogaster. Genetics 53, 237-248

The Mouse. (n.d.). Retrieved December 31, 2014, from http://web.jhu.edu/animalcare/procedures/mouse.html

The Rat. (n.d.). Retrieved December 31, 2014, from http://web.jhu.edu/animalcare/procedures/rat.html

Vanderlip, Sharon (2003). The Guinea Pig Handbook. Hauppauge, NY: Barron's

Von Frisch, K. (2011). Decoding the language of the bee. Readings in Zoosemiotics, 8, 141.

Walpole, S.C., Prieto-Merino, D., Edwards, P., Cleland, J., Stevens, G., & Roberts, I. (2012). The weight of nations: an estimation of adult human biomass. BMC Public Health, 12 (1), 439.

Washburn D.A. & Rumbaugh D.M. (1991) Ordinal Judgments if Numerical Symbols by Macaques (Macaca Mulatta). Psychological Science, 2 (3), 190-193

Watanabe, S. (2013). Preference for and Discrimination of Paintings by Mice. PloS one, 8 (6), e65335

White J.G., Southgate E., Thomson J.N., & S. Brenner (1986) The Structure of the Nervous System of the Nematode Caenorhabditis elegans. Philosophical Transactions of the Royal Society B: Biological Sciences. 314, 1-340

The validity of the Cattel-Horn-Carroll model on the intraindividual approach

Cristiano Mauro Assis Gomes[1], Jhonys de Araújo[1], Michele Gomes Ferreira[1] and Hudson F. Golino[2]

1 Universidade Federal de Minas Gerais, Brazil
2 Faculdade Independente do Nordeste, Bahia, Brazil

ABSTRACT

The Cattell-Horn-Carroll (CHC) model is considered the state-of-the-art of the psychometric tradition about intelligence. However, researchers of the dynamic systems field argue that the interindividual variation applied by psychometrics on intelligence field can produce inferences about the population but not about an individual. The present study investigated the validity of the CHC model at the level of the individual through the intraindividual approach. A dynamic factor analysis was employed in order to identify the factor structure of one individual scores on nine tests of the Higher-Order Cognitive Factor Battery, throughout 90 measurement occasions. Those tests measure, in the population level, three second order abilities and the general factor of the CHC model. Only the general intelligence factor was identified. Ultimately, the CHC model did not present validity to the assessed person. Implications for intelligence theories and measurement are discussed.

KEYWORDS: intelligence; CHC model, intraindividual approach, ergodic theorems

THERE IS A CONFLICT between two ways to study the human being and they are explained by the nomothetic and idiographic approaches. Whereas the idiographic approach tries to study singularity and particularities, the nomothetic approach studies groups and uses statistics and quantitative analysis to make generalizations. It can be observed that in the "fight" between these two approaches, the nomothetic approach is winning and has been the mainstream in psychology. One of the reasons for this is because the idiographic approach is many times seen as exclusively a clinical area, distant from the rigor and precision required by the science. However, there are studies, such as the one from Robinson (2011), that deny this vision and show that the idiographic versus nomothetic debate was misinterpreted. As pointed:

> "Idiographic knowledge aims at describing and explaining particular phenomena …. Nomothetic knowledge, on the other hand, has the aim of finding generalities that are common to a class of particulars and deriving theories or laws to account for these generalities" (Robinson, 2011, p. 1).

As stated by Robinson, the idiographic approach comes from the Wundtian idea that science needs to develop methodologies that adequately understand individual cases, or particular phenomena. So, in this sense it drives theorizations.

On the other side, it also serves as a way to test the theories, following the refutation processes that were formally proposed by Popper (1972). As Robinson's study points, Allport made a mistake when he equated the idiographic approach with a specific research method, i.e. the study of individuals, and the nomothetic approach with another specific research method: the study of groups and populations. In this way, the idiographic approach could be viewed as an enemy of science because it would be concerned only about singularity and particularities. However, this is not true, and Allport's dichotomization was a misconception of Windelband's concepts (Robinson, 2011). If the nomothetic approach aims for generality, this is only possible through the idiographic approach. The latter provides what Popper (1972) called "negative singular cases" providing to the former the refutability that are essential in science. It is only from the idiographic approach that we can see the "black swan" and refute the theory that was developed considering the other approach. Therefore, these two approaches should be seen as complementary and not opposite.

Many of the traditionally used methodologies in psychology, such as exploratory and confirmatory factor analysis, come from the general linear model that usually estimates groups and not indi-
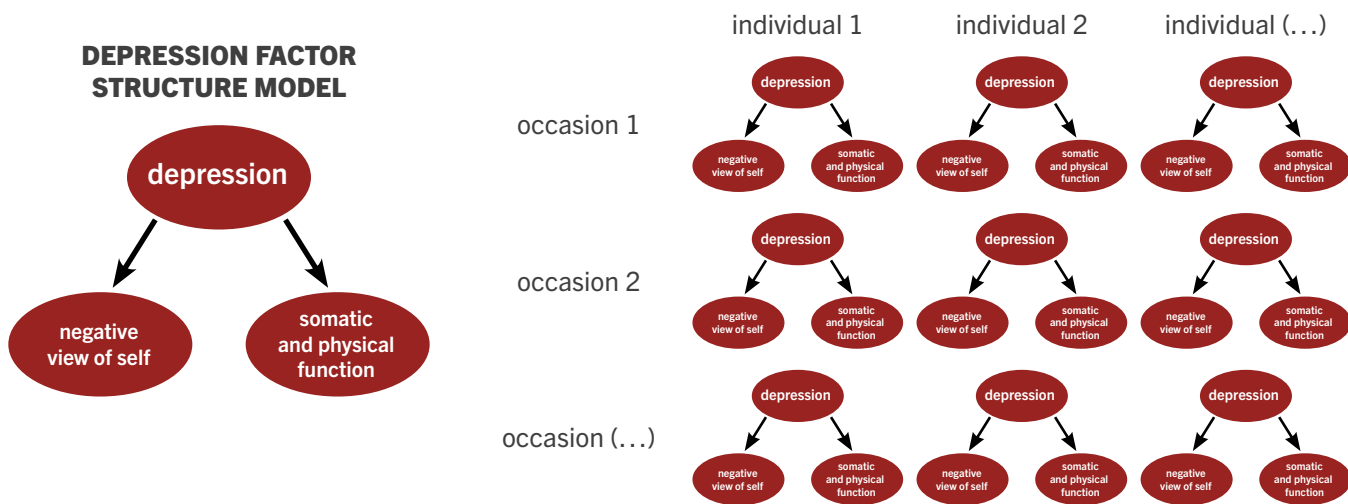
Figure 1. example of the depression factor model and the representation of the homogeneity and ergodicity conditions.

viduals. These methodologies are used to investigate the structure, organization and distribution of psychological constructs in the population level. They are tools from the nomothetic approach, called interindividual approach because their statistics are based in the differences between individuals (Molenaar, 2007A). Although providing important information for the understanding of psychological functioning in the population, the interindividual approach normally cannot provide information about how individuals function (Molenaar, Sinclair, Rovine, Ram & Corneal, 2009). The inappropriateness to directly transpose population information to individual was mathematically proved through the ergodic theorems.

The ergodic theorems, which come from the mathematical theory of ergodicity, are mathematical-statistical models that were first developed in the 1930's to study dynamic systems, due mainly to the work of Henri Poincaré, George Birkhoff and John Von Neumann (Ugalde, 2007) The ergodic theorems define two necessary and sufficient conditions that allow the generalization of knowledge from the interindividual structure observed in the population to the individuals: homogeneity and stationarity. The homogeneity criterion requires that all the individuals from the population have the same statistical structure and parameters (Moleenar, 2007). According to the stationarity criterion, this statistical structure cannot vary throughout the time. For example, consider a specific construct, such as depression. Suppose that it is composed by one higher-order factor (depression) and two specific factors: negative view of the self, and somatic and physical function (Figure 1). In order to follow the homogeneity criterion, all the individuals from the population should have the same factor structure and should follow the stationarity criteria. This factor structure could not vary across multiple measurement occasions. Most of the psychological processes violate both conditions, and therefore are considered non-ergodic processes. For this reason, it would not be possible to transpose the population data directly to the individual. Taking all this into consideration, Moleenar proposes an intraindividual approach.

The behavior and psychological characteristics of a person are usually non-ergodic (Nesselroade & Molenaar, 2010). Despite the "bad news", both intraindividual and interindividual variation can be analyzed at the same time. Raymond Cattell (1952), in the early 50´s, had already proposed a model to study the psychological variables that involved the analysis of the differences between individuals, called *R*-technique (interindividual), as well as the analysis of a single subject throughout the time, called *P*-technique (intraindividual) as can be seen in Figure 2. He explains that all the experimental designs in psychology have three components— individuals, time and variables, and depending on how they are combined, a different technique is used. The *R*-technique measures one or more variables, in several individuals, during a single occasion or a few occasions and allows identifying common factors in the population. The *P*-technique, by the other side, measures one or more variables in a single subject, during several occasions. However, the technique that
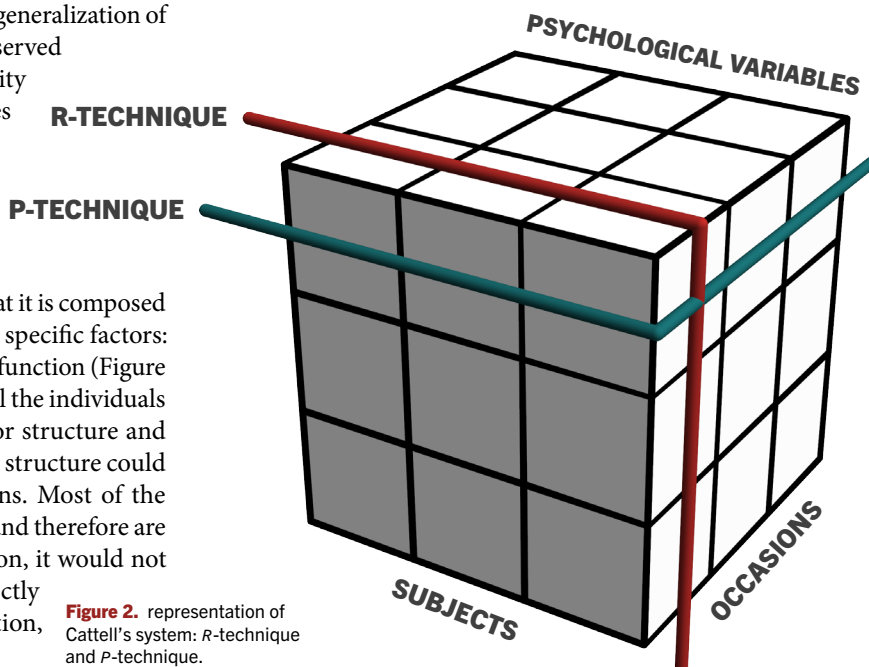
Figure 2. representation of Cattell's system: *R*-technique and *P*-technique.

Cattell proposed for an intraindividual analysis was the same used for the interindividual analysis: the *P*-factor analysis and the traditional factor analysis do not consider the score dependency across the time.

The *P*-technique was criticized because it is very similar to the traditional factor analysis and does not consider the score dependency across the time. If a person responds to an instrument 100 times, a correlation matrix will be produced from the 100 raw scores of the person data. That correlation matrix is similar to the traditional matrix used on factor analysis. It considers that the answers are independent and that the previous behaviors do not influence the next, which is rarely true (Ram, Brose, & Molenaar, 2013). Theoretically, previous answers predict or are related to the next and the *P*-technique does not incorporate this prerogative. An approach that aggregates the conditionality between the achievements is the dynamic factor analysis or the time series factor analysis. These involve a correlation matrix plus a time series analysis. Different strategies of data arrange and estimation are capable to adequately account for the relationships between responses over time. One of these is the Toeplitz matrix, for example.

Ram, Brose and Molenaar (2013) synthetize the main differences between the Cattell's *P*-technique factor model, time series factor model and the dynamic factor model in a very comprehensive way. The authors state that the objective of the *P*-technique is to describe relations among multiple responses of *P*-data, i.e. data collected in multiple occasions in one or more variables, in order to discover the structure underlying the responses or to test hypothesis regarding the day-to-day variation observed. However, as pointed before, since repeated measurements obtained from the same person are generally related, a key assumption required by factor analysis will probably be violated: the independency of the observations (Ram, Brose, & Molenaar, 2013). Ram, Brose and Molenaar (2013) point that in the years following the development of the *P*-technique factor model, a number of alternatives emerged to account for the relationship between the variables, for example the autoregression and moving average time series' models. In 1985 Peter Molenaar introduced the dynamic factor analysis as an alternative to *P*-technique factor model and to the time series models, since it enables two things: "deal with the independence violations and provide a framework for modeling the dynamic nature of ongoing processes" (Ram, Brose, & Molenaar, 2013, p.3). In the dynamic factor model, the multivariate state of an individual at any time is given by concurrent influences and past states (Ram, Brose, & Molenaar, 2013).

The psychometric approach to intelligence

In the early twentieth century, Spearman (1904) introduced a new perspective in psychology. He developed a key instrument to analyze data in psychology, called factor analysis, enabling the empirical research of psychological constructs through analysis of a correlation matrix on cognitive tests. Spearman (1904) identified one factor that explained the common variance among all IQ tests. Such factor was called *g* or general intelligence factor (Spearman, 1904). The specific variance not explained by *g* received the name of specific factor or *s*.

A body of studies was conducted throughout the years and divergent evidences were found. (Spearrit, 1996). In order to solve the problem of divergent evidences, John Carroll (1993) published in the early 1990's a meta-analysis study that included the main researches about intelligence conducted in the last 80 years. Carroll's findings suggested that intelligence has three levels or strata. The higher stratum, level 3, is composed by the general intelligence factor. The intermediate stratum, level 2, consists of eight broad cognitive abilities, while the basic stratum, level 1, has more than 50 specialized abilities.

In the late 1990's, McGrew and Flanagan (1998) proposed the integration of the Cattell-Horn-Carroll models, creating the CHC (Cattell-Horn-Carroll) model. This model consists of a multidimensional view of the intelligence, with three cognitive levels: the general factor (3rd level), 10 broad cognitive abilities (2nd level) and more than 70 specialized abilities (1st level). The CHC model has been validated in several papers, all around the world. In the Brazilian literature, for example, Gomes and Borges (2007), Gomes (2010) and Wechsler and Schelini (2006), found evidence supporting the CHC model. However, all of the validity studies of the CHC use the interindividual approach. Taking the ergodic theorems as reference, it is not possible to state that the three levels of the cognitive architecture found in the population are also present at the individual level without validation at the individual level.

The present study aims to analyze the validity of the CHC model in one individual, using the intraindividual approach. In order to do that, nine intelligence tests from the Higher-Order Cognitive Factor Battery (BAFACALO—Gomes, 2010) were administered to a single subject, on 90 different occasions. Using the time series and single case study design, we seek to verify if the CHC model is valid to explain the intraindividual variation of the scores from one individual over time. We expect to find, through the use of dynamic factor analysis, three latent variables (fluid intelligence—Gf, processing speed—Gs and crystalized intelligence—Gc) and at least a moderate correlation among them, indicating the presence of the general intelligence factor (*g*). In relation to the data of this study, the CHC model predicts that a general intelligence factor must be encountered. This general factor must explain three lower-order factors (fluid intelligence—Gf, processing speed—Gs, and crystalized intelligence—Gf). These three specific factors should explain, respectively, its specific marker tests. As pointed, the CHC model defines a hierarchical relationship between the cognitive abilities and it is expected that the commented hierarchy will be present. On the contrary, the CHC model could be refuted.

» METHOD
Participant

The subject of this study was a 23 year old student, who graduated in biological sciences at the *Pontifícia Universidade Católica do Paraná, Brazil*. At the time he participated in the study, he was studying ergodic theorems, but he did not have any knowledge about the psychometric models of intelligence or any previous contact with the tests used in this study. The participant is one of the co-authors of this study, but great care was taken to prevent his previous knowledge from affecting his performance in the

tests, creating bias in the results. The participant had not had any contact with the tests, before they were administered to him during the trial. Furthermore, at the moment that the participant responded the tests, he did not know anything about the psychometric intelligence models and the question of this study. He did all the tests in a blind way.

Instruments

Nine tests of the Higher-Order Cognitive Factor Battery were used. The Battery of Higher-Order Cognitive Factors or the BAFACALO's project, was developed by Gomes (2005, 2010), and was based on the Educational Testing Service's Kit of Factor-Referenced Cognitive Tests (1976). In the theoretical domain, the BAFACALO battery was developed to assess the general intelligence factor (*g*), plus six broad abilities presented in Carroll's three stratum theory (Carroll, 1993) and the CHC model (McGrew, Keith, Flanagan, & Vanderwood, 1997) in high school students. These broad abilities are: fluid intelligence (Gf) (Gomes & Borges, 2009A), crystalized intelligence (Gc) (Gomes, 2012), short-term memory (Gsm) (Gomes, 2011), broad visual perception (Gv) (2009B), fluency (Gr) and broad cognitive speediness (Gs).

Evidences show that the Higher-Order Cognitive Factor Battery is able to measure *g* and six broad cognitive abilities of the second level from the CHC model (McGrew et al., 1997). The structure of the battery was investigated by Gomes (2010) using Exploratory Factor Analysis (EFA) and Structural Equation Modeling (SEM). Each factor retention criterion used resulted in a different factor structure. The parallel analysis identified three factors, while the Kaiser criterion suggested a four factor structure. Both, the scree plot and the maximum likelihood approach, suggested a six factor structure. The second-order general factor was identified in every solution by a second-order EFA. Each structure pointed by the EFA result was tested via SEM. The three broad factors' model with a second order general factor presented the worst fit (χ^2/gl = 3.02, CFI = .87, RMSEA = .08). The four broad factors' model with a second order general factor presented a χ^2/gl of 2.45, a CFI of .91 and a RMSEA of .07, while the six broad factors' model with a second order *g* presented the best fit (χ^2/gl = 1.39, CFI = .98, RMSEA = .04).

Despite the sample of the original study presenting a broad variety of socio economic status (SES) levels, which is present in the Brazilian population of high school students, the sample was not intended to be representative of the Brazilian high school students. On the contrary, the sample is a convenient sample composed by 292 Brazilian high-school students from one public school (53.40% girls, 46.60% boys) of Belo Horizonte, Minas Gerais, Brazil. The majority of the sample was composed by girls, reflecting the demographic characteristic of the Brazilian population. Their age ranged from 14 to 20 years old (Mean *M* = 15.71, Standard Deviation *SD* = 1.15). Most of the participants had a monthly household income varying from R$1,751 to R$3,500 Reais. In order to recruit participants for this study, the school principal sent an invitation letter to all the students of the school, with the research purpose, the name and contact of the research team, as well as the dates of data collection. The chief of the school's Psychology Department visited every class reinforcing the Principal's invitation to participate in the study, and answered every question raised by the students.

Those interested in being part of the study were contacted by the researchers and signed a consent form, and confirmed to be at the school in the scheduled testing days. From the 320 students enrolled in the school, 91.25% accepted being part of the study, and answered the tests. It was not possible to know about sampling bias from the 8.75% of students that did not accept to participate of the study. The school only disposed data to the researchers about the students that accepted to participate of the research and signed the consent form. It is worth mentioning that Ethics Committees in Brazil does not allow incentives in researches involving the human being, so no incentive was given to the students.

All tests have Cronbach's alpha above .70, and also present structural, divergent, convergent, predictive and incremental validity (Gomes & Borges, 2009A, 2009B; Gomes, 2010; Gomes, 2011; Gomes, 2012). The first three tests listed measure the fluid intelligence (Gf), the three following tests measure the crystalized intelligence (Gc) and the last three tests measure the processing speed (Gs).

Inductive test (I). It consists of 12 items with an execution time limit of 14 minutes. Each item is composed by five groups of 4 letters. Among the 5 groups there are four groups that represent the same pattern, i.e. the letters are organized according to the same rule. The respondent has to identify the group of letters that does not follow the pattern and mark it with an ×.

Logical reasoning test (LR): This test consists of 30 items and the time limit for its execution is 24 minutes. Each item consists of a conclusion based on two abstract logical premises, with no relationship to the real world. The respondent has to indicate if the logical conclusion is appropriate or inappropriate.

General reasoning test (GR): It consists of 15 items and the time limit for its execution is 18 minutes. Each item is composed by a mathematical-logical problem, with a statement and a space to solve it. The respondent has to interpret the statement, solve the problem and choose one of the five multiple-choice answers.

Verbal comprehension test 1 (V_1): It consists of 24 items and the time limit for its execution is 6 minutes. Each item is composed of one reference word and five multiple-choice words. The respondent must identify the word which best approximates, in terms of meaning, to the reference word and mark it with an X.

Verbal comprehension test 2 (V_2): It consists of 18 items and the time limit for its execution is five minutes. Each item is composed of one reference word and five multiple-choice words. The respondent must identify the word which best approximates, in terms of meaning, to the reference word and mark it with an X.

Verbal comprehension test 3 (V_3): It consists of 18 items and the time limit for its execution is five minutes. Each item is composed of one word of reference and four multiple-choice words. The respondent must identify the word which best approximates, in terms of meaning, to the reference word and mark it with an X.

Perceptual speed test (P_1): The test consists of 410 words, whose 50 words begin with the letter "*a*" and the time limit is two minutes. The task is to mark all the words with the letter "*a*" in the given time Perceptual Speed Test 2 (P_2). The test has 48 items and its time limit is one and a half minute. Each item has the following task: to compare pairs composed by several digits and to identify if they are equal or different.

Perceptual speed test 2 (P₂): The test contains 48 items and its time limit is one minute and a half. Each item has a template figure and five options; only one option is identical to the template. The respondent must identify the figure that corresponds to the template.

Procedures

During approximately three months, the same nine intelligence tests from the Higher-Order Cognitive Factor Battery (BAFACALO) were administered to the same subject in 90 different occasions. The participant had contact with the tests only at the moment of the administration. After the last administration moment, the tests were scored by raw score sum. The score in each test corresponded to the number of correct answers of the participant. The scores were registered in Excel spreadsheet and each line corresponded to the score of the participant in a specific moment.

Data analysis

In temporal series, the performance is influenced by the previous performance. For that reason, the traditional exploratory factor analysis methods are not suitable to analyze this kind of data, as it assumes score errors' independence. The software DYFA2.03 from Browne and Zhang (2005) was used to analyze the scores produced during the 90 moments of the test administration. This software performs the dynamic factor analysis, what is appropriate to investigate data from temporal series. The dynamic factor analysis used the following basic equations (Ram, Brose, & Molenaar, 2013):

$$y_t = \Lambda \eta(t) + \varepsilon(t) \quad (1)$$

$$\eta(t) = B_1 \eta(t-1) + B_2 \eta(t-2) + \cdots + B_s \eta(t-s) + \zeta(t), \quad (2)$$

where y_t is a vector of the observable variables indexed by time ($t = 1, 2, \ldots, T$), Λ is the $p \times q$ factor loading matrix, $\eta(t)$ is a q-variate time series of latent factor scores and $\varepsilon(t)$ is the specific error plus measurement error time series. In equation 2, The $\eta(t)$ is modeled as a function of prior weighted (B_1 to B_s) latent states from $\eta(t-1)$ to $\eta(t-s)$. As pointed by Ram, Brose and Molenaar (2013), "present time "disturbances" are then introduced as a q-variate set of latent "innovations," $\zeta(t)$, and residual (measurement + specific) errors, $\varepsilon(t)$, the latter of which may be correlated across occasions." The common factors follow the VARMA process (p, q), generating a manifest-stationary temporal series. According to Browne and Zhang (2005), there is no likelihood function available for lagged covariance matrix, therefore the goodness of fit indexes are not applicable. The estimation method used is a simple discrepancy function of the minimum ordinary squares.

DYFA provides a series of rotation procedures of the factor solutions. The rotation used in this study was the Crawford-quartimax Ferguson, equivalent to quartimin, a solution that enables the identified oblique factors to correlate. The dynamic exploratory factor analysis used in this study was VARMA (1, 1) with a three

maximum lag, the number of occasions was 90, the number of individuals and common factors was one. The estimation approach used was the two stages method and the observable variables were the tests P_1, P_1, P_2, P_3, GR, LR, I, V_1, V_2, V_3, described previously.

» RESULTS

The participant's performance showed variance in all the tests, which is an essential condition for the intraindividual variation analysis. Table 1 shows the means, standard-deviations and the maximum performance in the nine tests employed in this study. The P_3 test was the easiest test, because the average of correct answers was 99.5%. The GR test was the second easiest. The most difficult test was P_2, since the average achievement was 53.58%. The test with the smallest performance variance was P_3, with a standard deviation of 2.06%. On the other hand, the P_1 test was the one that showed the highest performance variance, with a standard-deviation of 16.72%. There was a ceiling effect, where all items were passed by the participant, on the P_1, P_3, and GR tests.

The temporal series of the participant's performance during the 90 measuring occasions are shown on Figure 3. The P_3 test showed a very high growth rate between the first and second occasion. After the second administration moment the performance remained relatively constant, reaching an asymptote around the twentieth measuring occasion. The test I showed a similar pattern.

The V_1 and V_2 tests showed a high initial growth rate with some small decreases during that time. After that, a plateau-type performance is observed, followed by a sudden increase in the performance and a new plateau. The V_3 test showed a similar pattern to the V_1 and V_2 tests. However, the decreases present in the initial growth are more substantial. There is also a stronger change in the performance after the first plateau. The GR test shows a pattern relatively similar to the one found in V_2. However, GR does not present a plateau type pattern followed by a sudden change. In GR it is possible to observe an asymptote around the twentieth occasion. The performance in P_2 test was variable during all the time, with improvements and decreases over the whole period. Overall, there was a gradual increase in performance. The performance in P_1 showed an increase more constant than in the test P_2, appearing to be more gradual and characteristic of a logistic growth curve. The performance in LR showed a relatively erratic pattern, similar to the test P_2, at the beginning of the time series. However, the performance growth became more stable and gradual from the twenty-fifth occasion of measurement, and in the fiftieth occasion, it reached the asymptote.

Table 2 shows the correlations between the observable variables during lag zero. As can be seen, in general, the correlations are moderate to high suggesting the presence of a general factor. A superficial examination did not seem to indicate that the specific groupings P_1, P_2 and P_3 (group of tests to measure Gs), LR, I and GR (group of tests to measure Gf) and V_1, V_2 and V_3 (group of tests to measure Gc) have intragroup correlations higher than the

Table 1. means, standard-deviations and maximum of the observable variables

tests	M (%)	SD (%)	max (%)
P_1	87.26%	16.72%	100.00%
P_2	53.58%	6.94%	68.75%
P_3	99.51%	2.06%	100.00%
V_1	81.38%	6.83%	91.67%
V_2	73.33%	5.56%	83.33%
V_3	85.00%	5.39%	94.44%
GR	95.78%	11.65%	100.00%
LR	80.93%	9.81%	90.00%
I	89.91%	5.34%	91.67%
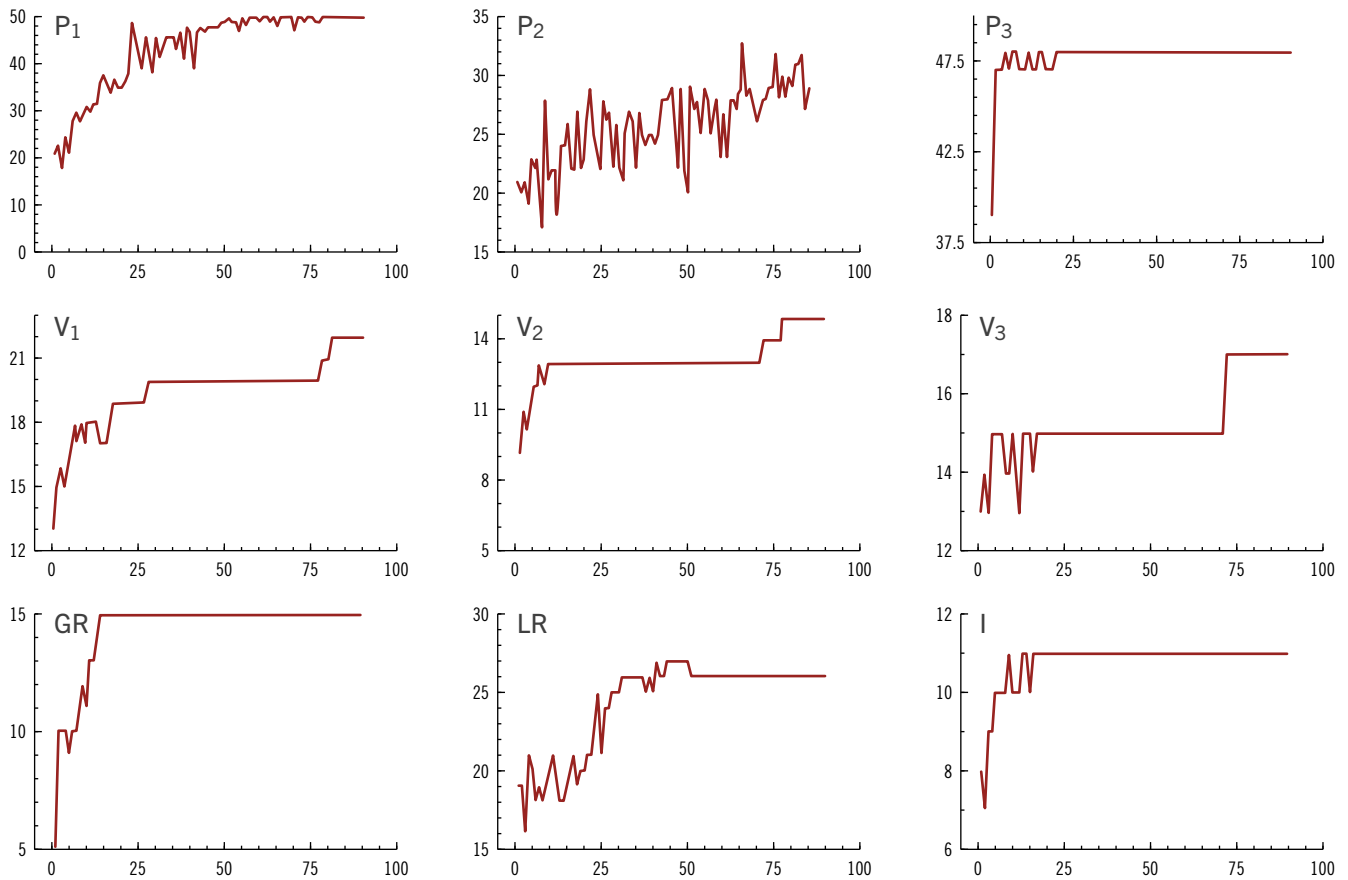
Figure 3. temporal series of the participant's performance in the 90 measuring occasions

extra-group correlations, possibly indicating that the three broad skills, Gf, Gs and Gc from the CHC model, are not identified in the individual examined.

The Cattell's scree test showed that a factor solution with a single factor was adequate. Figure 4 shows the eigenvalues in decreasing order and a line indicating which eigenvalues can be considered noise and should not be retained. This line crosses the second eigenvalue up to the sixth eigenvalue and it shows that only the first eigenvalue should be retained.

An exploratory dynamic factor analysis, with VARMA (1, 1) process, was performed using three lags to estimate the solution.

Table 2. correlation matrix of the observable variables during the lag zero

	P_1	P_2	P_3	V_1	V_2	V_3	GR	LR	I
P_1	**1.00**	**0.64**	**0.50**	0.87	0.66	0.59	0.80	0.89	0.73
P_2	**0.64**	**1.00**	**0.29**	0.65	0.62	0.57	0.48	0.57	0.48
P_3	**0.50**	**0.29**	**1.00**	0.59	0.54	0.34	0.69	0.42	0.59
V_1	0.87	0.65	0.59	**1.00**	**0.84**	**0.69**	0.78	0.80	0.76
V_2	0.66	0.62	0.54	**0.84**	**1.00**	**0.86**	0.64	0.51	0.64
V_3	0.59	0.57	0.34	**0.69**	**0.86**	**1.00**	0.43	0.49	0.42
GR	0.80	0.48	0.69	0.78	0.64	0.43	**1.00**	**0.65**	**0.85**
LR	0.89	0.57	0.42	0.80	0.51	0.49	**0.65**	**1.00**	**0.58**
I	0.73	0.48	0.59	0.76	0.64	0.42	**0.85**	**0.58**	**1.00**

In order to have a satisfactory exploratory dynamic factor solution, two conditions should be met. The first one is the identification of the VARMA model and the second one is the stationarity condition. The software DYFA 2.03 has two indicators, one for each condition. The solution of one general factor met both: stationarity of 0.91, smaller than 1 (necessary limit) and the average VARMA moving weight of 0.09, smaller than 1 (necessary limit).

Figure 5 shows how much the participant's performance in each test is correlated to his previous performance in the same test. This indicator is given by the ACF (auto-correlation factor) and it can be observed that, in the majority of the tests, the auto-correlation decrease as the intervals (lags) between one measurement and another measurement progress. This fact indicates that the previous performance is affected by the immediately previous performance and that the effect of this influence decreases along the occasions. Figure 5 also allows us to conclude that the choice for three lags was sufficient to identify the correlation matrix. In the axes of each test in Figure 5, there is a pair of horizontal lines bellow and above the x-axis, indicating how many lags are required for the correlation matrix to be captured by the factor solution. The PACF (*partial auto-correlation factor)* allows such identification, as it indicates if the auto-correlations residue is statistically significant according to the lags chosen. None of the partial auto-correlations is statistically significant after the third lag.
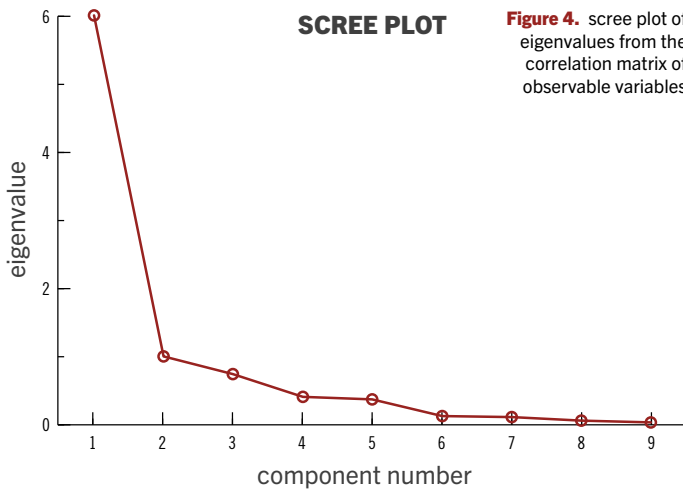
SCREE PLOT

(lag 1) of the general factor are highly correlated (.89). In turn, two measurement occasions of the general factor with another measurement time point between them (lag 2) also show a high correlation (.81).

» DISCUSSION

The aim of this study was to investigate the validity of the CHC model to identify the cognitive architecture of a specific individual, given the limitations of using population data to make inferences about individuals (Molenaar, 2007A, 2007B). So, the intraindividual approach analyzed the variance in the cognitive performance of one subject over 90 measurement occasions. The results are contrary to the presence of the expected three broad abilities (Gf, Gs and Gc). A single latent variable was identified indicating that the Spearman's model (1904) is the model which best explains the participant's performance, by the general factor (g) and the specific factor (s) of each test administered. The factor loadings found in this study varied from moderate-high to high, indicating that the general factor is of great importance to explain the variance of all tests administered. Another considerable part is explained by a specific factor in each test (Spearman's s factor), which is, at the same time, mixed with a portion of the measurement error.

The factor loadings of the tests in relation to the general factor were from moderate-high to high ($V_1 = .97$; $P_1 = .92$; GR = .85; $V_2 = .84$; $I = .80$; LR = .78; $V_3 = .68$; $P_2 = .67$; $P_3 = .61$, in descending order). The correlation between two measurement occasions of the general factor decrease progressively as the lags of such measurements increase. Two consecutive measuring occasions
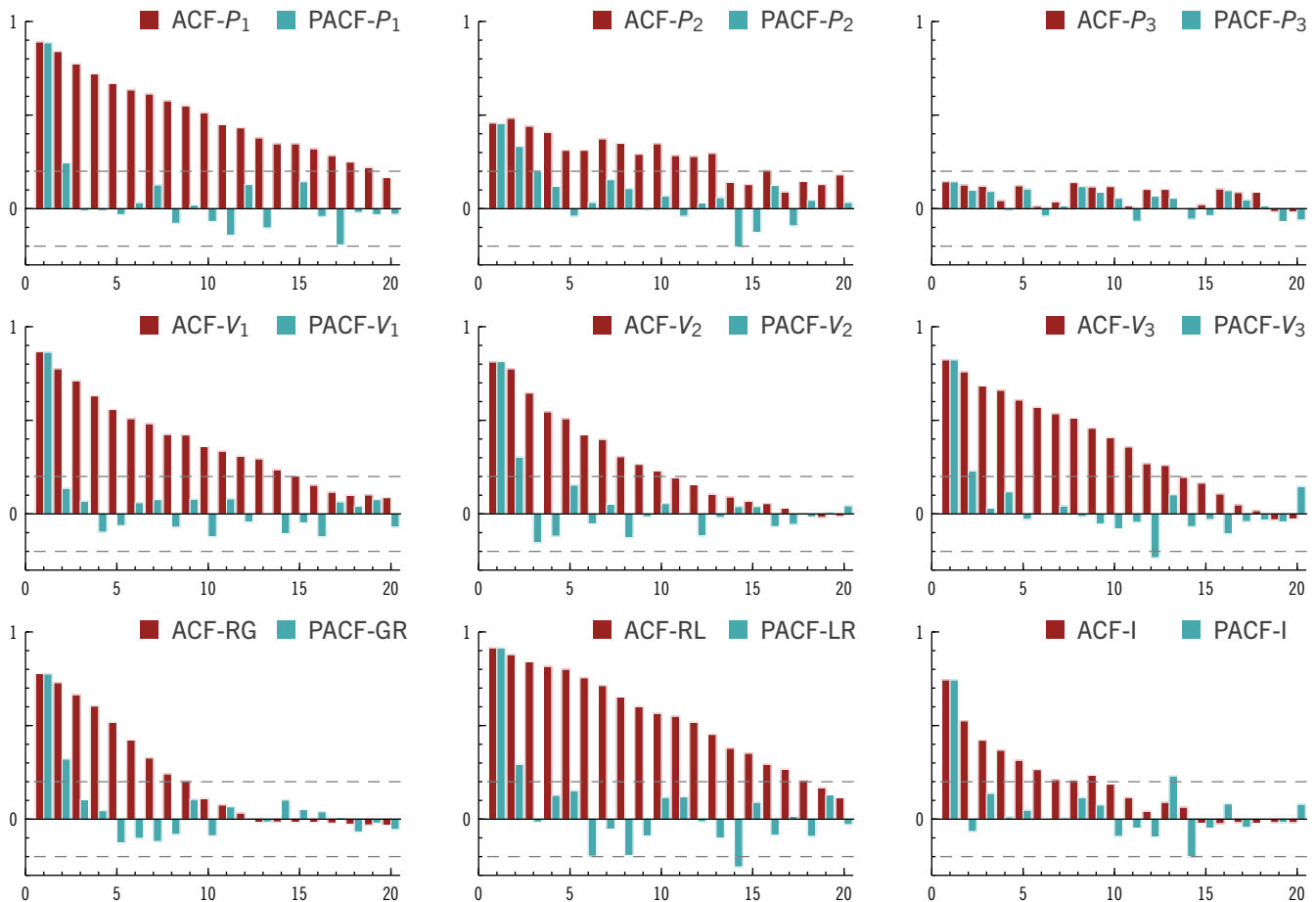
Figure 5. description of the auto-correlation factor (ACF) and partial auto-correlation factor (PACF) for each observable variable

In summary, Spearman's model (1904), the oldest one and the precursor of the factor studies in intelligence, is the model that explains properly the cognitive structure found in this study. It is interesting to note that when the intraindividual approach was used in a single case study, it returns to the model where it all began: Spearman's model.

The selection criterion of factor retention by the scree test can be questioned, since this criterion was essential for the evidence of a single factor. If another criterion was used, it could detect the presence of the three factors; in this case the CHC could be valid to explain the participant's performance. Two other criteria (eigenvalue greater than one and theoretical criterion) are commonly used to retain factors in an exploratory factor analysis. The first one determines the retention of factors that have an eigenvalue greater than one, being one of the most used criterion and a default procedure in popular statistical software, as the SPSS. Although being largely used, studies show that this criterion is inappropriate for small samples, i.e. less than 1,000 cases. The present study used a sample of 90 measurement occasions; therefore this criterion would not be appropriate. The second criterion usually employed is the theoretical one, in which the number of factors retained depends of the quantity of factors postulated theoretically. In this criterion, the appropriate number would be a solution with 3 factors.

Considering the possible argument that this study has not identified Gc, Gf and Gs only because of the retention criterion used, an exploratory factor analysis with three factors was conducted. The results did not indicate the presence of any of the broad abilities (Gf, Gc or Gs). Also, the three factors found in this particular analysis cannot be explained by any of the existing intelligence theories and models. Table 3 shows the tests factor loadings, by each factor. Factor 1 is strongly loaded by GR, I and P_3. Factor 2 is strongly loaded by V_2 and V_3 and factor 3 is strongly loaded by LR and P_1. The subject performance improved during the occasions and this may explain the factors. The tests LR and P_1 follow a similar path. Both indicate a gradual growth, but unstable, until half of the measurement occasions, when an asymptote is reached. However, P_2 loads moderately the factor 3 and does not represent this growth tendency. The test V_2 and V_3, which strongly loads factor 2, presents plateaus followed by a sudden change. The same pattern is shown in test V_1, but it present only moderately loads on factor 2. On the other hand, the P_2 test loads moderately the factor 2, but does not present the same growth tendency. The tests that better load factor 1 reach a quick growth, followed by a stable asymptote, until the end of the 90th measurement occasion. In sum, the three factors found can be interpreted as representing different growth trends.

However, it is not consistent with any traditional models of the interindividual approach and does not point to the validity of the CHC model, or the presence of Gf, Gs and Gc. The three factors presented a correlation range of .54 and .65, clearly indicating the occurrence of a general factor that has already been identified in the single factor solution.

Table 3. factors and factor loadings

	factor 1	factor 2	factor 3
P_1	0.20	0.08	**0.78**
P_2	−0.06	0.40	0.45
P_3	**0.71**	0.11	−0.10
V_1	0.31	0.39	0.43
V_2	0.23	**0.91**	−0.09
V_3	−0.17	**0.89**	0.15
GR	**0.91**	−0.03	0.13
LR	0.02	−0.03	**0.94**
I	**0.74**	0.07	0.13

To conclude, the single factor solution seems to be satisfactory to explain the participant's performance. All the tests present a good loading in the general factor, meaning they are well explained by the general factor. The general factor alone explains 67.28% of the participant's performance variance, while the solution with three factors explains 86.98% of the variance. Regarding the three factor solution, even if it had been chosen because of the theoretical criteria, it does not support the CHC model. However, the three factor solution brings new possibilities for future studies, taking into consideration the growth pattern of the participant's performance throughout the measurement occasions. New studies can investigate the empirical viability of these factors using a larger number of measurement occasions.

The results of the present study also suggest that population data cannot be used to make inference about single individuals' intelligence structure, a claim made several times both theoretically and empirically by Molenaar (2007A, 2007B). So we endorse that psychometric studies should expand their spectrum and not only invest in population based researches, but also in individual-based ones. Population and individual characteristics can coincide in some cases, but this does not seem to be the general rule. In order to better understand the complexity of the psychological processes and promote progress in the fields of Psychology and Psychometrics, it will be important to understand the convergences and divergences of these two approaches.

Future intraindividual studies should consider increasing the number of participants. The sample of this study was a single individual, who was a young adult, with normal development and with a University degree. Besides this limitation, the current study opens doors for an approach that is not commonly used in the psychometric literature on intelligence, and is used in a small scale in the Psychological literature worldwide. Future studies, with larger and more diverse samples, will be able to contribute for a better understanding about the different patterns of the individuals' cognitive architecture even though we understand that there may be some important obstacles to the intraindividual approach study. This type of study requires multiple measurement occasions and the availability of people to participate in such studies might be limited. A potential way to address this limitation is to use new technologies of data collection, such as smartphones and other web-related technologies. One good example of using new technologies to collect a large amount of data from individuals is the *Flu Near You* project (see: https://flunearyou.org/), which uses a smartphone's app to collect flu-related symptoms in a day-by-day fashion. In the same line, Van de Leemput et al. (2014) studied time series (200 assessment occasions) of four emotions in a large group of healthy and depressed people through an app on a smartphone. The authors point that nowadays is easier to assess and monitor psychological variables, such as mood indicators, in individuals due to the advancement of web applications (Van de Leemput et al., 2014). They provide a user-friendly interface in a tool (smartphone) that is used several times per day, and also enables the implementation of assessment feedback which may

increase the motivation of the participant to answer the research questions several times. Finally, another outstanding example of massive intraindividual data collection was provided by Myin-Germeys, Oorschot, Collip, Lataster, Delespaul and van Os (2009).

The current study used the intraindividual approach to evaluate a cognitive model and the instrument used was a particular battery.

Further studies could be conducted throughout the observation of specific behaviors, for example, during a certain period of time, as the above mentioned papers. Researchers that focus in the behavior analysis have already the expertise to observe and register behaviors and the intraindividual approach can bring them a contribution in relation to the data analysis. ■

....

REFERENCES

Browne, M. W. & Zhang, G. (2005). DyFA: Dynamic Factor Analysis of Lagged Correlation Matrices, Version 2.03 [Computer software and manual]. Retrieved from http://quantrm2.psy.ohio-state.edu/browne/ .

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

Cattell, R. B. (1952). *P*-technique factorization and the determination of individual dynamics structure. Journal of Clinical Psychology, 8(1), 5-10.

Ekstrom, R. B., French, J. W., Harman, H. H. & Dirmen, D. (1976). Manual for kit of factor-referenced cognitive tests. Princeton: Educational Testing Service.

Gomes, C. M. A. (2005). Uma análise dos fatores cognitivos mensurados pelo Exame Nacional do Ensino Médio (ENEM). Unpublished doctoral dissertation, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.

Gomes, C. M. A. (2010). Estrutura fatorial da bateria de fatores cognitivos de alta-ordem (BaFaCalo). Avaliação Psicológica, 9(3), 449-359.

Gomes, C. M. A. (2011). Validade do conjunto de testes da habilidade de memória de curto-prazo (CTMC). Estudos de Psicologia, 16(3), 235-242.

Gomes, C. M. A. (2012). Validade de Construto do Conjunto de Testes de Inteligência Cristalizada (CTIC) da Bateria de Fatores Cognitivos de Alta-Ordem (BaFaCAIO). Gerais: Revista Interinstitucional de Psicologia, 5, 294-316.

Gomes, C. M. A., & Borges, O. N. (2007). Validação do modelo de inteligência de Carroll em uma amostra brasileira. Avaliação Psicológica, 6(2), 167-179.

Gomes, C. M. A., & Borges, O. (2009a). Qualidades psicométricas do conjunto de testes de inteligência fluida. Avaliação psicológica, 8(1), 17-32.

Gomes, C. M. A., & Borges, O. N. (2009b). Propriedades psicométricas do conjunto de testes da habilidade visuo espacial. Psico-USF, 14(1), 19-34.

McGrew, K. S., & Flanagan, D. P. (1998). The intelligence test desk reference (ITDR) — Gf–Gc cross battery assessment. Boston: Allyn and Bacon.

McGrew, K.S., Keith, T.Z., Flanagan, D.P., & Vanderwood, M. (1997). Beyond g: the impact of Gf-Gc specific cognitive ability research on the future use and interpretation of intelligence tests in the schools. School Psychology Review, 26, 189-201.

Molenaar, P. C. M. (2007a). On the Implications of the Classical ergodic theorems: analysis of developmental process has to focus on intraindividual variation. Development Psychobiology, 50(1), 60-69.

Molenaar, P. C. M., Sinclair, K. O., Rovine, M. J., Ram, N., & Corneal, E. (2009). Analyzing developmental processes on an individual level using nonstationari Time Series Modeling. Developmental Psychology, 45(1), 260-271.

Molenaar, P. M. (2007b). Psychological methodology will change profoundly due to the necessity to focus on intra-individual variation. Integrative Psychological & Behavioral Science, 41(1), 35-40.

Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & van Os, J. (2009). Experience sampling research in psychopathology: opening the black box of daily life. Psychological Medicine, 39, 1533—47.

Nesselroade, J. R., & Molenaar, P. C. M. (2010). Analyzing intraperson variation hybridizing the ACE model with p-technique factor analysis and the idiographic filter. Behaviour Genetics, 40(6), 776-783.

Spearman, C. (1904). General intelligence, objectively determined and measured. American Journal of Psychology, 15(2), 201-293.

Ram, N., Brose, A., & Molenaar, P. M. (2013). Dynamic factor analysis: Modeling person-specific process. In T. D. Little (Ed.), The Oxford handbook of quantitative methods (Vol 2): Statistical analysis (pp. 441-457). New York, NY, US: Oxford University Press.

Spearrit, D. (1996). Carroll's model of cognitive abilities: educational implications. International Journal of Educational Research. 25(2), 107-198.

Thurstone, L. L. (1938). Primary mental abilities. Chicago: University of Chicago Press.

Van de Leemput, I. A., Wichers, M., Cramer, A. O. J., Borsboom, D., Tuerlinckx, F., Kuppens, P., & Scheffer, M. (2014). Critical slowing down as early warning for the onset and termination of depression. Proceedings of the National Academy of Sciences. Doi: 10.1073/pnas.1312114110.

Ugalde, E. (2007). De la mecánica estadística a la teoría ergódica. Revista Mexicana de Física, 53(2), 191-194.

Wechsler, S. M., & Schelini, P. W. (2006). Bateria de habilidades cognitivas Woodcock-Johnson III: validade de construto. Psicologia: Teoria e Pesquisa, 22, 287-295.

The small effects of non-hierarchical complexity variables on performance

Michael Lamport Commons[1], Sagun Giri[2], and William Joseph Harrigan[3]

1 Harvard Medical School
2 Dare Institute
3 Harvard University

Even when the results show that most of the difficulty in solving problems is explained by the hierarchical complexity of the item, there are still variables that help in small ways in predicting how well items measure difficulty. One must understand what these variables are and take them into consideration when analyzing data from instruments designed to measure the impact of the order of hierarchical complexity of items.

This study was designed to test the effect of small variables on task performance. The variables tested were hierarchical complexity, place in order, the number of calculations needs, the size of the numbers, and the causal variable position.

Participants were asked to solve problems from task sequences from the logic/mathematics/physical science subdomains. The four instruments used were the algebra, balance beam, infinity and laundry instruments. These instruments were based on the model of hierarchical complexity (MHC). Participants were asked to first complete the laundry task sequence and move to the next task sequence. Items from each instrument were analyzed individually and as a group. A Rasch analysis was performed on all the items from all the instruments. The variables thought to have an effect were coded. The coded variables were then analyzed using stepwise regression.

A stepwise regression was used and the small variables were tested with and without hierarchical complexity as a factor. The variables were regressed against the stage score of the items. For all four instruments stepwise regression with hierarchical complexity as one of the variable accounted for about 95% of the variance and the β was greater than 0.9. Stepwise regression with all the other variables except hierarchical complexity accounted for relatively lower variance and β.

The results showed that order of hierarchical complexity has a very strong predictive role and accounts for most of the variance. The other variables only made very small contributions.

KEYWORDS: model of hierarchical complexity, difficulty of items, small effect variables

EVEN WHEN THE RESULTS show that most of the difficulty in solving problems is explained by the hierarchical complexity of the item, there are still variables that help in small ways in predicting how well items measure difficulty. One must understand what these variables are and take them into consideration when analyzing data from instruments designed to measure order the impact of the order of hierarchical complexity of items. Some of these variables change depending on the instrument.

Results from the one domain paper showed that problem solving difficulty of instruments from logic/mathematics/physical sciences subdomain is explained by the order of hierarchical complexity of the item. Other instrument variables may help better predict how well items measure difficulty. These variables should be taken into consideration when analyzing data. This study was designed to test the effect of small variables on task performance. The variables tested were hierarchical complexity, place in order, the number of calculations needed, the size of the numbers, and the causal variable position.

Correspondence regarding this article should be directed to Michael Lamport Commons, Ph.D., Assistant Clinical Professor, Department of Psychiatry, Harvard Medical School, Beth Israel Deaconess Medical Center, 234 Huron Avenue, Cambridge, MA 02138-1328, commons@tiac.net.

» THE MODEL OF HIERARCHICAL COMPLEXITY

The four instruments used in the study were from logic/mathematics/physical sciences subdomain. The four instruments used were the algebra, balance beam, infinity and laundry instruments. These instruments were based on the model of hierarchical complexity. The model of hierarchical complexity (MHC) is a non mentalistic, neo Piagetian mathematical model (Krantz, Luce, Suppes, & Tversky, 1971; Luce & Tukey, 1964). MHC allows for the measurement of stage performance. It deconstructs tasks into the actions that must be done at each order. This is to build the behavior needed to successfully complete a task.

MHC provides an analytic *a priori* measurement of the difficulty of task actions. The difficulty is represented by the orders of hierarchical complexity (OHC) (Commons & Pekker, 2008). There are 17 known orders of hierarchical complexity. This is shown in Table 1

Hierarchical complexity describes a form of information that is different from traditional information theory (Shannon & Weaver, 1948) in which information is coded as bits that increase quantitatively with the amount of information. Theorem 4 of the model (Commons, et al, 1998) shows that every task action has an order of hierarchical complexity associated with it. The ideal correct task actions may be classified as to their order of hierarchical complexity. The tasks actions may address every experimental task, every clinical test item that has a difficulty associated with it, every behavior, developmental task, survey item, and statement made by people regardless of the content or context. Each task action will have a difficulty of performance associated with it.

A task action is defined as more hierarchically complex when 1) A higher-order task is defined in terms of two or more tasks at the next lower order of hierarchical complexity, 2) Higher-order tasks organize the lower order actions and 3) The lower order tasks are coordinated non-arbitrarily, not just put together as an arbitrary chain. This is illustrated schematically in Figure 1.

Table 1. the 17 known orders of hierarchical complexity

order	
number	name
0	computational
1	automatic
2	sensory or motor
3	circular sensory motor
4	sensory-motor
5	nominal
6	sentential
7	preoperational
8	primary
9	concrete
10	abstract
11	formal
12	systematic
13	metasystematic
14	paradigmatic
15	crossparadigmatic
16	meta-crossparadigmatic

» THE RASCH MODEL

Whereas they are well-known in psychometric circles, Rasch's (1980), models for measurement have been employed by developmental psychologists only recently (Andrich & Constable, 1984; Bond, 1994; Dawson, 1998, 2000, 2002; Draney, 1996; Muller, Sokol, & Overton, 1999; Wilson, 1984). These models are designed specifically to examine hierarchies of person and item performance, displaying both person proficiency and item difficulty estimates along a single interval scale (logit scale) under a probabilistic function. In addition, they can be employed to test the extent to which items or scores conform to a theoretically specified hierarchical sequence. A central tenet of stage theory is that cognitive abilities develop in a specified sequence, making the statistical tests implemented in a Rasch analysis especially relevant to understanding stage data. The Rasch model permits

researchers to address questions like, "Are all single abstractions items more difficult than all representational systems level and less difficult than all abstract mappings items?" Moreover, the detailed information about item functioning and individual performances provided by the software makes it possible to simultaneously examine group and individual effects. These properties make Rasch models uniquely suitable for the investigation of many developmental phenomena.

The Rasch (1980) model uses logistic regression to minimize person and item error simultaneously. The model allows researchers to convert raw scores into equal interval linear scales. The item scores on the right represent how difficult the item was. The person scores on the left represent how good a person was at dealing with the item difficulty. The model also produces an objective, additive, and one dimensional scale. These are some of its advantages over other scaling techniques in measuring stage.

It is beyond the scope of this paper to provide a comprehensive account of the Rasch model, though we do attempt to provide enough information to allow readers who are unfamiliar with the model to follow the results of the analysis. For an introduction to the Rasch model, see Bond and Fox (2001), Rasch (1980), Smith (2004), and Wilson (2005).

» METHOD

Participants

The convenience sample of 309 participants was obtained from various ListServ's. There were 232 (77.3%) men and 77 (22.7%) women ranging in age from 12 to 87 ($M = 24.02$, $SD = 8.86$) with education varying from elementary school to graduate degree ($M = 4$, $SD = 1.01$)

Instruments

A number of task sequences were used. The task sequences were categorized as either belonging to the mathematical (algebra & infinity), logical (laundry version of the Inhelder and Piaget's (1958) pendulum problem), or physical science (balance beam) subdomains.

Task sequences from the mathematics, logic and physical science subdomains were used. Participants were asked to solve problems from a task sequence with a given content. There were a total of 250 items.

Table 2. number of participants for each instrument of the social subdomain and mathematical/logical/physical science subdomain

instrument	number of participants
laundry	109
algebra	37
balance beam	46
infinity	50

A higher order action is:

1) defined in terms of the task actions from the **next lower order** of hierarchical complexity.

2) The higher order task action organizes two or more next lower order of hierarchical complexity.

3) The ordering of the lower task actions have to be carried out **non-arbitrarily**.

| order n + 2 |
| action 1 |

| order n + 1 | | order n + 1 |
| action 1 | | action 2 |

| order n | order n | order n | order n |
| action 1 | action 2 | action 3 | action 4 |
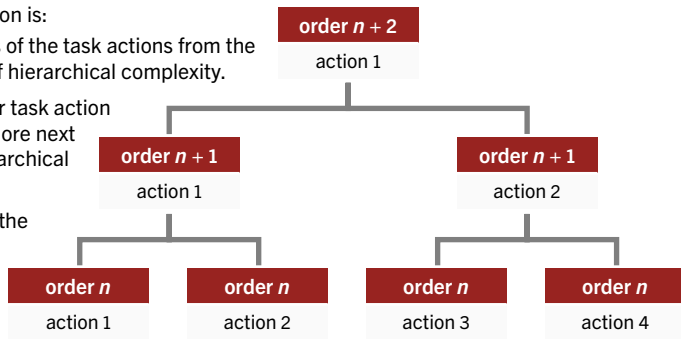
Figure 1. This figure schematically shows the three axioms of the model of hierarchical complexity

Each task sequence was at a different order of hierarchical complexity. These included:

1. Preoperational

2. Primary

3. Concrete

4. Abstract

5. Formal

6. Systematic

7. Metasystematic

Instrument description for logic/mathematics/ physical sciences subdomain

Algebra (mathematics): This instrument sequence asked participants solve standard algebraic problems

Balance beam (physics): This instrument sequence was derived from Inhelder and Piaget (1958) balance beam task

Infinity (mathematics): This instrument sequence asked participants to solve problems that dealt with conceptualizing infinity

Laundry (logic and mathematics): This instrument sequence was a version of Inhelder & Piaget (1958) pendulum task. It asked participants to detect causal relationships from various systems and then compare the systems using logic (Bernholt, Parchmann, & Commons, 2009; Commons, Miller, & Kuhn, 1982).

An example of logic/mathematics/physical sciences subdomain problem sequence is shown in Appendix A.

Procedure

Participants were asked to solve problems from a task sequences from the logic/mathematics/physical science subdomains. Participants were asked to first complete the laundry task sequence and move to the next task sequence. Items from each instrument were analyzed individually and as a group. A Rasch analysis was performed on all the items from all the instruments. The variables thought to have an effect were coded. The coded variables were then analyzed using stepwise regression.

Instrument variables affecting performance

Number of calculations required to solve the task: This was coded by counting each calculation required to solve the task and putting down the actual number. No shortcuts were taken so each number represented the maximum amount of steps a person had to take in order to correctly solve the task.

Size of numbers in the task: This was coded by counting the digits of the numbers in the problem and the answer choices. The number of digits in the largest number was used.

Where problem is within an order: This variable is determined by the order in which the problems were asked. For example, if in the primary order (8) there were five questions then the first question was coded as 1 and the last question was coded as 5. It is expected that the questions towards the end of an order would be easier because of the practice effect. With all the instruments, within each order of hierarchical complexity, an item presented became progressively more difficult as the regular complexity of the expressions representing a variable increased

Where the item is within a problem: Some of these variables change depending on the instrument. The last variable only occurred in the laundry instrument. It was used in place of the number size variable as the laundry instrument did now have numbers. This variable was coded based on the position of the causal variable. The variable which operates on the outcome is known as the operative variable. This is the variable which must be known in order to answer the question correctly. In the four variable laundry instrument the difficulty of the item changed depending on the position of the causal variable of the outcome within the episode. When first it was second easiest (primacy), when in the middle it was the most difficult and when right before the outcome the easiest (recency.)

What is the position of the causal variable: The variable which operates on the outcome is known as the operative variable. This is the variable which must be known in order to answer the question correctly. The laundry problem in Appendix B demonstrates this clearly.

Table 3. stepwise regression for all instruments *without* hierarchical complexity

model 1						
balance beam		**algebra**		**laundry**		**infinity**
size of numbers	$\beta = .906$ $p \le .001$ $n = 46$	place in order	$\beta = .366$ $p = .047$ $n = 37$	calc. needed	$\beta = .829$ $p \le .001$ $n = 109$	n/a
model 2						
balance beam		**algebra**		**laundry**		**infinity**
calc. needed added	$\beta = .926$ $\Delta R^2 = .037$ $n = 46$	size of numbers added	$\beta = .545$ $\Delta R^2 = .163$ $n = 37$	variable position added	$\beta = .855$ $\Delta R^2 \le .001$ $n = 109$	n/a
size of numbers	$\beta = .702$ $p \le .001$ $n = 42$	place in order	$\beta = .436$ $p = .013$ $n = 37$	calc. needed	$\beta = .739$ $p \le .001$ $n = 109$	n/a
calc. needed	$\beta = .279$ $p = .016$ $n = 46$	size of numbers	$\beta = -.410$ $p = .019$ $n = 37$	variable position	$\beta = .228$ $p \le .001$ $n = 109$	n/a

Table 4. stepwise regression for all instruments *with* hierarchical complexity

model 1							
balance beam		**algebra**		**laundry**		**infinity**	
OHC	$\beta = .977$ $p \le .001$ $n = 46$	OHC	$\beta = .953$ $p \le .001$ $n = 37$	OHC	$\beta = .964$ $p \le .001$ $n = 109$	OHC	$\beta = .912$ $p \le .001$ $n = 50$

model 2							
balance beam		**algebra**		**laundry**		**infinity**	
size of numbers added	$\beta = .991$ $\Delta R^2 = .003$ $n = 46$	calc. needed added	$\beta = .968$ $\Delta R^2 \le .001$ $n = 37$	calc. needed added	$\beta = .969$ $\Delta R^2 = .010$ $n = 109$	calc. needed added	$\beta = .969$ $\Delta R^2 = .010$ $n = 50$
OHC	$\beta = .735$ $p \le .001$ $n = 46$	OHC	$\beta = .943$ $p \le .001$ $n = 37$	OHC	$\beta = .830$ $p \le .001$ $n = 109$	OHC	$\beta = .830$ $p \le .001$ $n = 50$
size of numbers	$\beta = .128$ $p = .035$ $n = 46$	calc. needed	$\beta = .117$ $p = .024$ $n = 37$	calc. needed	$\beta = .168$ $p \le .001$ $n = 109$	calc. needed	$\beta = .168$ $p \le .001$ $n = 50$

» RESULTS

Variables combined

A stepwise regression was used and the small variables were tested with and without hierarchical complexity as a factor. The variables were regressed against the stage score of the items.

Stepwise regression without OHC

For the balance beam instrument without hierarchical complexity in the stepwise regression outputted two models. In the first model, size of numbers accounted for about 91% of the variance ($r(46) = .906, p \le .001$). In the second model size of numbers and calculations needed accounted for about 93% of the variance ($r(46) = .926, r^2$ change = .037). Size of numbers was responsible for about 70% ($r(46) = .702, p \le .001$) of the variance and calculations needed was responsible for about 28% of the variance ($r(46) = .279, p = .016$). Please refer to Table 3.

The algebra instrument without hierarchical complexity in the stepwise regression outputted two models. In the first model, the place of the problem within the order of hierarchical complexity accounted for about 37% of the variance ($r(37) = .366, p = .047$). The second added size of the numbers to the previous model. This addition accounted for about 55% of the variance ($r(37) = .545, r^2$ change = .163). The place in the order was responsible for about 44% ($r(37) = .436, p = .013$) of the variance and the size of the numbers was responsible for about 41% of the variance ($r(37) = −.410, p = .019$). Please refer to Table 3.

The laundry instrument without hierarchical complexity in the stepwise regression outputted two models. In the first model, the number of calculations required accounted for about 83% of the variance ($r(109) = .829, p \le .001$). The second model added the position of causal variable to the previous model. This addition accounted for about 86% of the variance ($r(109) = .855, r^2$ change = .044). The number of calculations needed was responsible for about 74% ($r(109) = .739, p \le .001$) of

the variance and the position of causal variable was responsible for about 23% of the variance ($r(109) = .228, p \le .001$). Please refer to Table 3.

The infinity instrument without hierarchical complexity included in the stepwise regression could not be computed because the effect of the small variables was not large enough.

Stepwise regression with OHC

For the balance beam instrument with hierarchical complexity as a variable in the stepwise regression, there were two models and the numbers change drastically. In first model, hierarchical complexity alone, accounted for about 98% of the variance ($r(46) = .982, p \le .001$). The second model added the size of numbers to the previous model. This addition accounted for just over 99% of the ($r(46) = .991, r^2$ change = .003). Hierarchical complexity accounted for about 74% ($r(46) = .735, p \le .001$) of the variance, and the size of the numbers accounted for about 13% of the variance ($r(46) = .128, p = .035$). Please refer to Table 4.

The algebra instrument with hierarchical complexity in the stepwise regression outputted two models. In the first model, hierarchical complexity alone accounted for about 95% of the variance ($r(37) = .953, p \le .001$). The second model added the number of calculations needed to the previous model. This addition accounted for about 97% of the variance ($r(37) = .968, r^2$ change = .013). Hierarchical complexity accounted for about 94% ($r(37) = .943, p \le .001$) of the variance and the number of calculations needed accounted for about 12% of the variance ($r(37) = .117, p = .024$). Please refer to Table 4.

The laundry instrument with hierarchical complexity in the stepwise regression outputted two models. In the first model, hierarchical complexity alone accounted for about 97% of the variance ($r(109) = .964, p \le .001$). The second model added the number of calculations to the previous model. This addition accounted for 97% of the variance ($r(109) = .969, r^2$ change = .010). Hierarchical complexity accounted for 83% ($r(109) = .830, p \le .001$) of the variance and the number of calculations required accounted for about 17% of the variance ($r(109) = .168, p \le .001$). Please refer to Table 4.

The infinity instrument with hierarchical complexity as a variable in the stepwise regression outputted two models. In the first model, hierarchical complexity alone accounted for 91% of the variance ($r(50) = .912, p \le .001$). The second model added the number of calculations needed to the previous model. This addition accounted for 97% of the variance ($r(50) = .969, r^2$ change = .010.) Hierarchical complexity accounted for 83% ($r(50) = .830, p \le .001$) of the variance and the number of calculations required accounted for about 17% of the variance ($r(50) = .168, p \le .001$). Please refer to Table 4.

Table 5. Pearson correlations with the major variables that affected performance

hierarchical complexity and stage score	$r = 0.955$
calculations required and hierarchical complexity	$r = 0.542$
calculations required and stage score	$r = 0.571$
calculations required and causal variable position	$r = 0.395$
causal variable and stage score	$r = 0.520$
causal variable and hierarchical complexity	$r = 0.527$

Table 6. stepwise regression for all instruments without hierarchical complexity

calculations needed	$\beta = .592$	$p \leq .001$
excluded variables		
place in order	$\beta = .094$	$p = .268$
size of numbers	$\beta = -.011$	$p = .894$

Table 7. stepwise regression for all instruments with hierarchical complexity

model 1		
hierarchical complexity	$\beta = .955$	$p \leq .001$
excluded variables		
place in order	$\beta = .020$	$p = .506$
calculations needed	$\beta = .075$	$p = .037$
size of numbers	$\beta = .029$	$p = .343$
model 2		
number of calculations needed added	$\beta = .957$	$\Delta R^2 = .004$
hierarchical complexity	$\beta = .914$	$p \leq .001$
calculations needed	$\beta = .075$	$p = .037$
excluded variables		
place in order	$\beta = .029$	$p = .339$
size of numbers	$\beta = .024$	$p = .422$

Collinearity

Some of the small variables correlated with the major variables that affected performance. The order of hierarchical complexity was correlated with the number of calculations required ($r = .542$) and the position of the causal variable ($r = .527$). The stage score of the items also correlated with the number of calculations required ($r = .571$) and the position of the causal variable ($r = .520$.) The number of calculations required and the position of the causal variable had a correlation of $r = .395$ and the order of hierarchical complexity and the stage score had a correlation of $r = .955$.

Overall regression analysis

An overall stepwise regression analysis was conducted across all instruments and the small variables were tested with and without hierarchical complexity as a factor. The variables were regressed against the stage score of the items. Without hierarchical complexity, the number of calculations accounted for 59% of the variance ($r(242) = .592, p \leq .001$). The place of the problem within the order of hierarchical complexity and size of numbers were excluded from the regression because of the effect of these variables were not large enough. Place in the order accounted for 9% of the variance ($r(242) = .094, p = .268$) and size of numbers accounted for 1.1% of the variance ($r(242) = -.011, p = .894$).

With hierarchical complexity as a variable in the stepwise regression there were two models. In the first model, hierarchical complexity alone as predictor accounted for about 96% of the variance ($r(242) = .955, p \leq .001$). In the first model, place of the problem within the order of hierarchical complexity, number of calculations needed and size of numbers were excluded from the regression because of the effect of these variables were not large enough. The second model added number of calculations needed to the previous model. This addition accounted for 96% of the variance ($r(242) = .957, r^2$ change $= .004$). Hierarchical complexity accounted for 91% ($r(242) = .914, p \leq .001$) of the variance and the number of calculations required accounted for about 7% of the variance ($r(242) = .075, p = .037$). In the second model, place of the problem within the order of hierarchical complexity, and size of numbers were excluded from the regression because of the effect of these variables were not large enough.

» DISCUSSION

Order of hierarchical complexity has a very strong predictive role and accounts for most of the variance. The other variables only made very small contributions. In the case of Balance Beam, the following small variables accounted for .991 Of the variance: hierarchical complexity, size of numbers in the task, number of calculations needed and place of the problem within the order of hierarchical complexity.

The small variables each would make a contribution when analyzed alone. When hierarchical complexity is used in the overall stepwise regression it wipes out almost all of the effect of the small variables. Number of calculations needed is the only variable with an effect. This variable had a β of only .075. The reason it had such a small effect was *a)* hierarchical complexity already accounted for most of the variance and *b)* calculations required was highly collinear with hierarchical complexity, $r = .542$. ∎

REFERENCES

Andrich, D., & Constable, E. (1984, March). *Studying unfolding developmental stage data based on Piagetian tasks with a formal probabilistic model*. Paper presented at the Annual meeting of the AERA, New Orleans.

Bond, T. G. (1994). Piaget and measurement II: Empirical validation of the Piagetian model. *Archives de Psychologie*, 63, 155–185.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement for the human sciences*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Commons, M. L., & Pekker, A. (2008). Presenting the formal theory of hierarchical complexity. *World Futures: Journal of General Evolution*, 64(5-7), 375-382. doi:10.1080/02604020802301204

Dawson, T. L. (1998). "*A good education is.*" A lifespan investigation of developmental and conceptual features of evaluative reasoning about education. Unpublished doctoral dissertation, University of California at Berkeley, Berkeley, CA.

Dawson, T. L. (2000). Moral reasoning and evaluative reasoning about the good life. *Journal of Applied Measurement*, 1(4), 372–397.

Dawson, T. L. (2002). A comparison of three developmental stage scoring systems. Journal of *Applied Measurement*, 3, 146–189.

Draney, K. L. (1996). *The polytomous Saltus model: A mixture model approach to the diagnosis of developmental differences*. Unpublished doctoral dissertation, University of California at Berkeley, Berkeley, CA.

Inhelder, B., & Piaget, J. (1958). *The Growth of Logical Thinking from Childhood to Adolescence*. Basic Books, New York, NY.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement*, Vol. I: Additive and polynomial representations. (pp. 93–112). New York: Academic Press.

Luce, R. D. & Tukey, J. W. (1964). Simultaneous conjoint measurement: a new scale type of fundamental measurement. *Journal of Mathematical Psychology*, 1(1), 1–27. doi:10.1016/0022-2496 (64)90015-X

Müller, U., Sokol, B., & Overton, W. F. (1999). Developmental sequences in class reasoning and propositional reasoning. *Journal of Experimental Child Psychology*, 74, 69–106.

Rasch, G. (1980). *Probabilistic model for some intelligence and attainment tests*. Chicago, IL: University of Chicago Press.

Shannon, C. E., & Weaver, Warren. (1963). *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.

Smith, R. M. (Ed.) (2004). *Introduction to Rasch measurement*. Maple Grove, MN: JAM Press

Wilson, M. (1984). *A psychometric model of hierarchical development*. Unpublished doctoral dissertation, University of Chicago, Chicago, IL, USA.

Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

APPENDIX

» APPENDIX A: *logic/mathematics/physical sciences subdomain problem sequence*

Balance beam metasystematic problem sequence example

Subsystem A: In one class, $x + 6$ students each received a rating of 1. The other 5 students received a rating of x.

ratings	1	x
number of students	x + 6	5

Subsystem B: In one class, $y + 6$ students each received a rating of 500. The other 5 students received a rating of $500 y$.

ratings	500	500y
number of students	y + 6	5

» APPENDIX B: *laundry problem*

Read the six examples. Each example tells what was done to the cloth and how the cloth turned out. Give the answer to the questions that follow the examples.

A cloth was stained with red lipstick. There are six ways the cloth can be washed. Sometimes the cloth will be clean after being washed and sometimes the cloth will be dirty.

A bleach	powder soap	blue booster	cold water	⟶	dirty
B bleach	liquid soap	pink booster	hot water	⟶	clean
A bleach	powder soap	pink booster	hot water	⟶	dirty
B bleach	powder soap	pink booster	cold water	⟶	dirty
A bleach	liquid soap	blue booster	hot water	⟶	clean
B bleach	liquid soap	blue booster	cold water	⟶	clean

Question

Look back at the examples. After being washed, will the cloth be clean or dirty?

				clean?	dirty?
B bleach	powder soap	blue booster	hot water		
A bleach	liquid soap	blue booster	cold water		
A bleach	powder soap	pink booster	cold water		
B bleach	liquid soap	blue booster	hot water		
B bleach	powder soap	blue booster	cold water		
B bleach	powder soap	pink booster	hot water		
A bleach	liquid soap	pink booster	hot water		
A bleach	powder soap	blue booster	hot water		
A bleach	liquid soap	pink booster	cold water		
B bleach	liquid soap	pink booster	cold water		
A bleach	liquid soap	pink booster	cold water		

BEHAVIORAL
DEVELOPMENT
BULLETIN

Advances in the model of hierarchical complexity (MHC)

Michael Lamport Commons[1] and Shuling Julie Chen[2]

1 Harvard Medical School
2 Harvard University

The model of hierarchical complexity (MHC) has a long history that only in very recent years resulted in its formal specification as a general model. A brief history of the origins of the notion of hierarchical complexity follows. This is done in order to identify what the steps were and also those who played key roles in the model's development over these many years.

In tracing the evolution of the MHC, there are four periods, each involving different people. The two earliest periods were Commons' pre-college years, college and graduate school years, followed by a period of more active and direct development from 1973 to 1982. After the model was developed, there were advances in the period from 1982 to present.

KEYWORDS: history, stage, model of hierarchical complexity, neo-Piagetian, mathematical developmental theory

» SOME BACKGROUND ON THE MODEL OF HIERARCHICAL COMPLEXITY

Pre-college years

When Commons was 12, he read Isaac Asimov's (1951) *Foundation.* The character Hari Seldon had a new theory of psychohistory about the mathematical, psychological and historical character of societal evolution. Today, it would be seen as a contribution to cultural evolution. The main idea was that mathematical psychology could be used to predict the future or at least understand it. His account was in narrative form.

College and graduate school years

While at Columbia University as a graduate student in the Faculty of Pure Science from 1966 to 1973, Joan Borrison, Commons' partner at the time, was taking a course from Jane-Ellen Huttonlocker and Brian Sutton-Smith at Teachers College, Columbia University. She was complaining bitterly about the book she was assigned. It was a book by Inhelder and Piaget (1958). The book turned out to be a life turning point for Commons. He found Inhelder and Piaget's account of development was fascinating. He felt strongly that there was something about Inhelder and Piaget theory that was right. Commons thought that it explained why people were not crazy per se or irrational in their actions. Rather, he thought human beings simplified the world in a fashion that resulted in illusions. But Commons did not agree with all of Inhelder and Piaget's ideas. He did not like the logical framework that Inhelder and Piaget employed in their model because it was based on a form of rationality (i.e. logic). Logic is too restrictive. Others

(Wason, 1968) have shown that the logical basis for their formal stage failed empirically. There are other relationships between events and actions besides logic. But, Commons thought Inhelder and Piaget had the right stages when the substages were counted as stages. Pascual-Leone (1970) has also successfully argued this case. To Commons, stage theory seemed to be much broader than Inhelder and Piaget had set forth. It also seemed that stage theory could be applied to all organisms' actions and not limited to just humans. The reliance on the "clinical method" that required interviews did not seem necessary.

1973 to 1982

Commons' realization was that task analysis could be used to replace the "clinical" method. Task analysis breaks down tasks into a sequence of task actions that are sufficient to accomplish that task. Task analysis comes from within behavior analysis, information processing theory, and psychophysics. With 12 years of mathematics, physics and experimental psychology as an undergraduate and graduate student, Commons thought about tasks in a very precise and mathematical way.

In 1975, while at *Northern Michigan University*, the idea of developing the model of hierarchical complexity started when Commons and Patrice Miller met Deanna Kuhn again at the meeting of the Society for Research in Child Development which met in Denver. She previously was at *Columbia University.* At the time she was a professor at *California State University at Fullerton*. They worked on some developmental research using the plant problem (Kuhn & Brannock, 1977; Kuhn, Ho & Adams, 1979). The plant problem is a rough analog of Inhelder and Piaget's pendulum problem (Inhelder & Piaget, 1958). In the plant problem, a participants' task

was to isolate the variable that predicted plants' health condition and exclude the variables that would not. The possible variables were: *a)* having or not having leaf lotion; *b)* dark or light plant food; *c)* a lot or little water. One of the differences between the plant problem and the pendulum problem is that the plant problem was something that people had much more experience with. Few people would have taken physics, which would have taught them about the pendulum.

Table 1. example of the plant problem's informational episodes (Commons, 1975)

big pot	leaf lotion	dark plant food	little water	⟶	sick
small pot	no leaf lotion	light plant food	lots of water	⟶	healthy
big pot	leaf lotion	light plant food	lots of water	⟶	sick
small pot	leaf lotion	light plant food	little water	⟶	sick
big pot	no leaf lotion	dark plant food	lots of water	⟶	healthy
small pot	no leaf lotion	dark plant food	little water	⟶	healthy

Table 2. Six out of the ten sample test questions: Look back at the examples. After being treated in the way shown, will the plant be healthy or sick?

small pot	leaf lotion	dark plant food	lots of water	⟶	healthy or sick?
big pot	no leaf lotion	light plant food	little of water	⟶	healthy or sick?
big pot	leaf lotion	dark plant food	little of water	⟶	healthy or sick?
small pot	leaf lotion	light plant food	lots of water	⟶	healthy or sick?
small pot	no leaf lotion	light plant food	lots of water	⟶	healthy or sick?
big pot	no leaf lotion	dark plant food	little water	⟶	healthy or sick?

Commons thought that for a well-designed psychology experiment, one had to better control how the variables and their values would appear in the task. The initial solution to this was to add a variable to Kuhn's version of the plant problem. The values for the variable that Commons added in 1976 were: Big Pot or Small Pot (See Table 1). In this new version, a plant was treated in six ways, as laid out in the first, six informational episodes. Sometimes it would be healthy after being treated this way and sometimes it would be sick. After reading the informational episodes, participants would then be shown the test episodes, which they were to respond to. Six of these are shown in Table 2.

As shown in Table 1, not only was it important that there be four variables that were always present in each episode, but there should be also only two values of each variable. Each variable had to occur in the same physical location; otherwise it would make it harder to see it as a clear variable and lead to confusion. There had to be exactly six episodes in the informational part of the task. With only four episodes, the task demands made the problem at the abstract order. There had to be two more plausible values of each possible cause (variables) per episode and two less plausible values of each possible cause per episode. All odd number versions would not have equal number of healthy or sick plants. With four variables and two values per variable, this produces $2^4 = 16$ possible forms for an episode.

Note that "cause" is really bidirectional here. One could predict the outcome from the "cause" and the outcome could predict the "cause." Last, the variables in an episode could occur in 4! = 24 orders. That meant, if variable order were changed, there could be 24 different problems.

Even with controlling for the constraints discussed above, no matter what was tried, there were issues with the new version of the plant problem that could not be solved. In that version, as in the simpler three variable version. all single cause episodes also have the complement set of the other variables as causes. So, if one variable, such as the amount of water given the plant, was to be associated with the plant being healthy, a combination of the rest would also be associated with the plant being healthy. The goal of only having a single cause turned out to be impossible.

This was confirmed empirically, after administering the new version to a group of participants, when there were two people who used a combination of variables to correctly predict the outcome. Commons realized that there were still problems that could not be solved even after all the balancing of all the values of the variables and outcomes. The mathematical fact was that there are always dependencies of causes in any finite set. Any finite causal system that has a single causal variable will also have the complement of the rest of the variables as "causal" variables as well. Only a system with infinite causal information could have only one cause.

That year, Commons, Miller, and Kuhn (1982) conducted a study of what courses people took if they scored concrete stage versus formal stage. The last problem of the plant problem was that the variable of whether leaf lotion was used or not had a "negative" versus a complementary value. All the other variables were complementary choices. This was corrected in the laundry problem which is used up until the present. An example of the laundry problem is shown in Table 3.

Table 3. Example of the laundry problem (1977): A cloth was stained with red lipstick. There are six ways it can be washed. Sometimes it will be clean after being washed and sometimes it will be dirty.

brand A bleach	powdered soap	blue booster	cold water	⟶	dirty
brand B bleach	liquid soap	pink booster	hot water	⟶	clean
brand A bleach	powdered soap	pink booster	hot water	⟶	dirty
brand B bleach	powdered soap	pink booster	cold water	⟶	dirty
brand A bleach	liquid soap	blue booster	hot water	⟶	clean
brand B bleach	liquid soap	blue booster	cold water	⟶	clean

Table 4. Six sample test questions of the possible ten of the laundry problem: Look back at the examples. After being washed, will the cloth be clean or dirty?

brand B bleach	powdered soap	blue booster	hot water	⟶	clean or dirty?
brand A bleach	liquid soap	pink booster	cold water	⟶	clean or dirty?
brand A bleach	powdered soap	blue booster	cold water	⟶	clean or dirty?
brand B bleach	powdered soap	pink booster	hot water	⟶	clean or dirty?
brand B bleach T	liquid soap	pink booster	hot water	⟶	clean or dirty?
brand A bleach	liquid soap	blue booster	cold water	⟶	clean or dirty?

The limitations of the Plant problem and Piaget's theory's led Commons (1976A, B) to view the constraints of each structural aspect of the problems. In comparing the two problems, Commons contrasted two properties of the problems: *a)* What was causal in the plant problem was arbitrary versus what was causal in the pendulum problem was fixed; *b)* Any finite causal system that had a single causal variable will also have the complement of the rest of the variables as "causal" variables. An infinite causal system could have only one cause. Commons then recognized that reflecting on the structure of the problems could not be done within the formal stage. He found that comparing the structural aspects within the problems formed two separate systems with different properties. One property was the arbitrariness of what would be a cause and the other was the finite versus infinite source of information about possible causes. Hence, he had identified a new stage, first called *the structural analytic stage* and later the metasystematic at the suggestion of Deanna Kuhn around 1980.

After coming to Harvard in the fall of 1977 as a postdoctoral fellow of Deanna Kuhn's, Commons began to work with Francis Asbury Richards (Rick Richards). They wanted to develop a problem that would take insights from Commons' comparisons of the differing structures of Kuhn and Brannock's (1977) plant problem and Inhelder and Piaget's (1958) pendulum problem. At that time, there was no adequate means to decide whether the structural analytic stage was a new stage or just another version of the formal stage. Commons and Richards constructed the four-story problem to show the difference between the structural analytic stage and the formal stage. The structure of the plant problem was still used, but they had variants with different content: the laundry, pendulum (Inhelder & Piaget, 1958) and paint versions. Richards and Commons developed the V. P. Vanktesh story, Bad Bart the gambler, the Washing machine story, and the Richard Regan story (Commons & Richards, 1978; Commons, Richards and Kuhn, 1982). The problem was a rough analog of Gödel's (See Heijenoort, Jean van, 1967) incompleteness problem in the sense that systems of formal stage relationships would be compared. It was at that point we divided the "stage demands" of the task from the stage of performance on that task, something that few other researchers considered.

In 1978, Commons and Richards began to develop their first pass at the *general stage model* which later became the *model of hierarchical complexity*. The model was later published in the book *Beyond Formal Operations* (1984). Identification of the metasystematic stage was one of the foundations of what led to the general stage model. Commons and Richards started with just three stages: *a)* the first being Inhelder and Piaget's (1958) formal stage; *b)* the second was the systematic stage that was suggested by Herb Koplowitz (1981). The case for this stage was later published in his chapter (Koplowitz, 1984). Koplowitz presented a case which he called cyclical causality; *c)* The third stage was the metasystematic (formerly called the structural analytic stage). The model was published as the general model of stage (Commons

A higher order action is:

1) defined in terms of the task actions from the **next lower order** of hierarchical complexity.

2) The higher order task action organizes two or more next lower order of hierarchical complexity.

3) The ordering of the lower task actions have to be carried out **non-arbitrarily**.
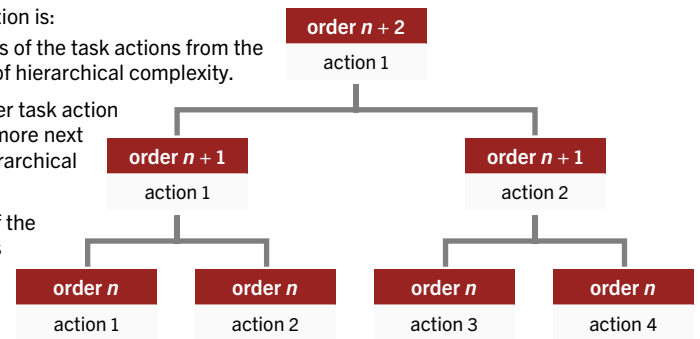
Figure 1. the three axioms of the MHC

& Richards, 1984). Theo Linda Dawson suggested the name the model of hierarchical complexity around 1997.

What is the model of hierarchical complexity?

There are two kinds of complexity of task actions. The commonly recognized one refers to the horizontal complexity often measured by the number of bits. Each "yes or no" question that is answered by correctly completing any task is representative of just one bit. For more horizontally complex tasks, the issue is how many bits of information there are embedded in a task. The number of bits is descriptive of the amount of horizontal complexity.

In contrast, the model of hierarchical complexity offers a different method of analyzing the difficulty of tasks. The model of hierarchical complexity (MHC) is a neo-Piagetian mathematical model. It is a newer form of stage theory that is both an advance on and a simplification of Piaget. MHC allows for the measurement of stage performance. It deconstructs tasks into the component actions that must be done to complete a task correctly at each order of hierarchical complexity (OHC). Each of these actions is necessary to build the behavior needed to successfully complete a task. A higher order action is: *a)* defined in terms of tasks actions from the *next lower* order of hierarchical complexity; *b)* the higher order task action *organizes* two or more less complex actions; *c)* the ordering of the lower task actions have to be carried out *non-arbitrary* way. This is shown in Figure 1.

The model of hierarchical complexity (MHC) identifies 17 *orders of hierarchical complexity*. Hence it looks at how the required task actions are organized into a "tree" structure (Figure 1). It is designed to account for development within a life-span and evolution across species. Stage is simply the order of hierarchically complexity of a task correctly carried out. The model's unidimensional *measure of order of hierarchical complexity* is on an equally spaced ordinal scale. The tasks are that of organizing actions that address problems presented or inferred. hierarchical complexity applies to any tasks and the set of events or occasions within them in which information is organized. The kinds of entities that organize information include humans and their biological systems as well as their social organizations, nonhuman organisms, and machines, including computers. The reason it applies so broadly is that it is a simple mathematical method of specifying tasks, and the tasks may contain any kind of information. Thus, its use of purely quantitative principles makes it universally applicable in any

context. This enables a standard quantitative analysis of hierarchical complexity in any setting, because it eliminates dependence on mentalistic, cultural, or other contextual explanations.

1982 to present

In 1981 and 1982, in many phone calls with Kurt W. Fischer, he suggested the name abstract stage for the early formal stage of Inhelder and Piaget's. He also suggested that the preoperational stage be split and that Commons read Biggs and Collis (1982). They had identified what we call the sentential stage 6 coming before the preoperational stage 7. The even-earlier preoperational stage that preceded the sentential stage was what Commons later called nominal stage 5. Fischer also suggested Commons and Richard read his seminal article (1980) in which Fischer laid out a good deal of the stage sequence. Commons and Richard were not aware of the article before, otherwise it would have saved them a lot of work. Yet they came up with almost the same stage sequences as Fischer's levels independently. Commons and Richards credited Fischer's work in the 1984 book but their model differed from what Fischer had. The general stage model had more stages: stage 0 for computers, stage 5 nominal, stage 14 paradigmatic, and stage 15 cross-paradigmatic. It would be flexible because it was possible to add stages in between, which happened in 2014 as described later. This was because there were no tiers that would fix the number of level to four per tier. Also there was no end to the stages as shown by the addition of stage 16 in 2013 as will also be discussed later.

The crucial insight that solved the major problem in Commons and Richard's 1984 version of the general stage model came when he was driving to Mexicali with Roger Dunn in 1984. Commons was working on collaborating with Jesús Francisco Galaz Fontes of Facultad de Ciencias Humanas de la Universidad Autónoma de Baja California, Mexicali. The collaboration was on the existence of formal stage reasoning in non-literates. He was explaining to Dunn about his model. Dunn asked how the organization of action was different from a chain of behavior. Commons answered that in a chain, the organization of the subtask actions were arbitrary, whereas in the organization of lower stage tasks the organization of actions could not be arbitrary. This is because that organization usually has to work in the real world. The conversation was crucial because they agreed on the nonarbitrary requirement of developing each stage. It was clear that higher order complexity task actions had to be defined in terms of lower order ones. This concept was published in Commons, Richards and Armon (1984).

As Commons, Richards and Armon (1984) were editing the *Beyond Formal Operations* book, they decided that it could initiate a new field of positive adult development .The new positive adult development field would be different from the traditional field of adult development and aging because the postformal stages from metasystematic on develop only from early adulthood on. Hence these developments were termed positive adult development. The stages of development in that field would be more coherent if they had a comparative table of stages from the different postformal stage proposers (Commons, Richards and Armon (1984).

More stages were added in the order in which they are discussed: *a)* stage 14 for the paradigmatic stage, which is a prerequisite for stage 15 cross-paradigmatic stage; *b)* computational stage 0 for computers and programming was added after the book; *c)* the meta cross-paradigmatic stage 16 was added in 2013; *d)* in 2014, the sensory or motor stage was split with the new automatic stage 1 being added.

In 1990, Jonalu Johnstone, Jeremy B. Straughn, Maryellen Meaney, Julia H. Weaver, Erica Lichtenbaum, Sharon R. Krause, with Dorothy L. Danaher, Cheryl Armon, Suzanne Benack, and Dawn Schrader wrote the first scoring manual called *Applying the General Stage Scoring System* (GSSS, 1990). This was to help other people to score interviews and later to construct instruments using the model of hierarchical complexity. The present system for constructing instruments based on systematically constructed vignettes began with Joseph Anthony Rodriguez (1992). Rodriguez and Commons adapted the stages for making the multisystems task to the new Doctor–Patient vignettes (Commons & Rodriguez, 1993). Theo Dawson made a key contribution in 1996 when she said Rasch analysis should be used to analyze results instead of signal detection.

R. Duncan Luce told Commons and his son Lucas in the summer of 2003 that distribution was the core idea that made things work in measurement theory. In mathematics, one of the differences between a ring and a group is that a ring has two mathematical operations with distributivity, for example times and plus. Here, Commons had generalized distributivity into the non-arbitrary organization of lower stage actions. This was general enough to fit all of thought and action, and yet powerful enough to generate stages. Luce mentioned that this can be clearly seen in the mathematical notion of distribution. The non-arbitrary axiom has major implications. For example, at the concrete stage 9, there is irreducibility of long multiplication to simple addition and multiplication.

The most extensive revision of the model of hierarchical complexity was completed in August of 2004. After working a year and a half with R. Duncan Luce, Alexander Pekker and Commons straightened out the formal mathematical theory. Pekker was a graduate student in mathematics at Harvard. He cleaned up the mathematics of the model. He also systematized the meaning of non-arbitrary ordering of lower order actions with combinatorial mathematical theory, showing that not all combinations of behaviors were allowed.

Later, in 2012, Commons' research assistant Eva Yujia Li mentioned that in order for a person to understand the crossparadigmatic stage, that person has to be at one stage higher. In this case, because commons is the person who invented crossparadigmatic stage, he himself had to be as least one stage higher. The invention can be done by induction. The higher stage was later named meta-crossparadigmatic stage.

Starting from July 2014, a new stage was developed. Before 2014, the model of hierarchical complexity (MHC) had 16 orders. However, applying the model to explain the development of operant conditioning (originally order 2) from respondent conditioning (originally order 1) in non-human animals, led to the discovery of another new stage. The old stage sensory or motor order 1 had to be separated into two new orders. The revised order 1 was renamed the *automatic* order 1 and the new order 2 as *sensory or motor*. Thus there are now 17 orders of hierarchical complexity and their corresponding stages.

» ADVANCES

From that early period, there have been a very large number of advances in the model of hierarchical complexity. The implications of the model have become much clearer across the years. These include understanding the nature of development as well as how when combined with the value of consequences of behavior, a very powerful model for predicting most behavior has emerged. The generality to all domains has also expanded. A sample of these advances are listed next with references when these references are available.

1. Commons, M. L., Richards, F. A. & Kuhn, D. (1982). Systematic and metasystematic reasoning: A case for levels of reasoning beyond Piaget's stage of formal operations, *Child Development,* 3, 1058–1069. Retrieved from http://dareassociation.org/Papers/Systematic%20and%20Metasystematic%20reasoning1982.pdf

This was the first publication to introduce four stages after Piaget's formal stage, and the general stage model which was later to be called the model of hierarchical complexity. Different from other attempts in the developmental literature to try to postulate "postformal operational" stage of cognition, the model of hierarchical complexity's four postformal stages constituted a quantitatively distinct structure or stage of action, problem solving and reasoning. An instrument (four story problem) was developed to assess these modes of cognition, in systematic and metasystematic stage reasoning. The first European scholar who used the same four story problem (Commons, Richards & Kuhn, 1982) was Eeva Kallio. She published her first article based on her MA thesis in developmental psychology (Kallio & Helkama, 1991). She continued critical discussion of these three stages (Kallio, 1995).

2. Commons, M. L., Miller, P. M. & Kuhn, D. (1982). The relation between formal operational reasoning and academic course selection and performance among college freshmen and sophomores. *Journal of Applied Developmental Psychology,* 3, 1–10. Retrieved from http://www.dareassociation.org/Papers/The%20Relation%20between%20formal%20operational%20reasoning1982.pdf

This study showed that concrete stage performing students did not take any courses with mathematics in them including all the science courses. Formal stage performers according to the MHC and Inhelder and Piaget's criteria did. This showed that stage was a strong predictor of other behaviors, such as academic course selection.

3. Commons, M. L. & Rodriguez. J. A. (1990). "Equal access" without "establishing" religion: the necessity for assessing social perspective-taking skills and institutional atmosphere. *Developmental Review,* 10, 323–340. Retrieved from http://dareassociation.org/Papers/Equal%20Access%20without%20Establishing%20religion1990.pdf

This paper introduced the application of the model of hierarchical complexity to stages of social perspective taking. It showed

how MHC could be applied to major legal decisions. There is a possibility that the Supreme Court or other courts have used this article in some of their decisions.

4. Commons, M. L. (1991). A comparison and synthesis of Kohlberg's cognitivedevelopmental and Gewirtz's learning-developmental attachment theories. *J. L. Gewirtz & W. M. Kurtines (Eds.) Intersections with attachment,* 257–291). Hillsdale, NJ: Erlbaum Associates, Inc.

Kohlberg's cognitive developmental and Gewirtz's learning developmental approaches to attachment were compared. The two approaches were synthesized yielding the sequence of attachment stages and stage-change processes.

5. Commons, M. L., Krause, S. R., Fayer, G. A., & Meaney, M. (1993). Atmosphere and stage development in the workplace. In J. Demick & P. M. Miller (Eds.). *Development in the Workplace* (pp. 199–220). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. Retrieved from http://dareassociation.org/Papers/CommonsKrauseFayerMeaney1993.pdf

This paper presented a method for characterizing the relationship between individuals and their workplace environment with respect to individual moral development. The model of hierarchical complexity was used to investigate and characterize individual development and the stage of development embodied in the workplace environment.

6. Sonnert, G., & Commons, M. L. (1994). Society and the highest stages of moral development. *Politics and the Individual,* 4 (1), 31–55. Retreived from http://dareassociation.org/Papers/Society%20and%20Highest%20Moral%20Dev.html

This paper discussed the higher stages in Kohlberg's stage theory of moral development. It resulted in the first re-definition of Kohlberg's highest moral stages 5 (metasystematic) and 6 (paradigmatic) in a non-arbitrary way that satisfied the criteria from the model of hierarchical complexity for what constitutes a stage.

7. Commons. M. L., Lee. P., Gutheil, T. G., Rubin, E., Goldman, M., & Appelbaum, P. L. (1995) Moral stage of reasoning and the misperceived "duty" to report past crimes. *International Journal of Law and Psychiatry,* 18 (4), 415–424. Retrieved from http://dareassociation.org/Papers/Moral%20State%20of%20Reasoning1995.pdf

This paper examined clinicians' sensitivity to the rights of patients in two domains: the false duty to report past crimes (misprision) and the duty to report patients' future potential for violence. This was the first application of the model of hierarchical complexity to the professional relationship between doctors and their patients.

8. Commons, M. L., & Goodheart, E. A. (1999). The philosophical legacy of behaviorism. *The Philosophical Foundations of Behaviorism,* 9–49.

The paper argues that as decentration progresses in human beings, mentalistic notions of causes of behavior such as free will are replaced by non-mentalistic or more behavioral notions of causes. Behavior analysis was explained with respect to the model of hierarchical complexity. In the model of hierarchical complexity, an event is processed as data and classified in terms of task-required hierarchical organization of required response. By this time, the name model of hierarchical complexity had superseded the general stage model.

9. Commons, M. L., & Wolfsont, C. A. (2002). A complete theory of empathy must consider stage changes. *Behavioral and Brain Sciences*, *25* (1), 30–31.

In the paper, a sequential, hierarchical stage model of empathy accounted for a comprehensive range of empathic behaviors. Empathy is part of emotions. This paper showed that the model accounted for emotional development and not just cognitive development. In the paper, an illustrative table, "Stages of Empathy" was provided to demonstrate how increasingly hierarchically complex empathic behaviors emerge at each stage, beginning with the infant's "automatic empathy" and ending with the advanced adult's "coconstruction of empathetic reality." It showed that the model of hierarchical complexity can characterize empathy behaviors from infants to adulthood.

10. Commons, M. L., & Richards, F. A. (2002). Organizing components into combinations: how stage transition works. *Journal of Adult Development*, *9* (3), 159–177.

This was the first paper to explicate the nature of transition steps between stages in the model of hierarchical complexity. The model of hierarchical complexity of tasks leads to a quantal notion of stage, and therefore delineates the nature of stage transition. The model shows that there was only one stage sequence for stage transition.

11. Commons, M. L., & Richards. F. A. (2003). Four post formal stages the handbook of adult development, *The Springer Series in Adult Development and Aging*, 199–219. Retrieved from http://www.tiac.net/~commons/Four%20Postformal%20Stages.html

A detailed introduction of the four postformal stages was presented in this paper. The term "postformal" has come to refer to various stage characterizations of behavior that successfully address more hierarchically complex behaviors than those behaviors found in Piaget's last stage—formal operations. These stages are generally seen only in adults. Other postformal research that was directed toward an understanding of development in a single domain. The model of hierarchical complexity is standard universal model that addresses all tasks in all domains.

12. Commons, M. L., & White, M. S. (2003). A complete theory of tests for a theory of mind must consider hierarchical complexity and stage. *Behavioral and Brain Sciences, 26* (5), 606–607.

This was the first presentation to introduce stacked neural networks by using the model of hierarchical complexity as a frame. It explained the limitation of the traditional neural networks and the advances of the new hierarchical stacked neural networks.

13. LaLlave, J. (2005). The acceptability of arguments in favor of and against the Iraq War. *Conflict & Communication Online*, *4* (2), 1–16. Retrieved from http://dareassociation.org/Papers/Iraq%20War.LaLlave.pdf

This paper presented research pertinent toward the prevention and constructive transformation of conflicts. It proposed models that have external validity in the fields of political argumentation expert evaluation and testimony, psychological assessment and journalist reporting.

14. Commons, M. L., Rodriguez, J. A., Adams, K. M., Goodheart, E. A., Gutheil, T. G., & Cyr, E. D. (2006). Informed consent: do you know it when you see it? Evaluating the adequacy of patient consent and the value of a lawsuit. Psychiatric Annals, 36, 430–435. Retrieved from http://dareassociation.org/Papers/Informed%20Consent%202006.pdf

In this paper, the helper-person instrument was used to test how well practitioners engaged in informed consent. The helper person instrument consisted of six vignettes ranging in hierarchical complexity from primary to metasystematic. This measurement of individuals' stage of reasoning about helping predicted how positively mental health practitioners are seen and whether or not they will be sued and for how much.

15. Ross. S. (2006). effects of A structured public issues discourse method on the complexity of citizens' reasoning and local political development (doctoral dissertation). Retrieved from http://dareassociation.org/Papers/Sara%20Nora%20Ross%20Dissertation.doc

The model for hierarchical complexity was used to structure the reasoning about issues that would increase in political development. This shows that the model is good for accounting for phenomena in the separate social science of political science, and not just simple behaviors.

16. Robinett, T. (2006). Moral reasoning and political affiliation in liberal and conservative voters: applying a model of hierarchical complexity (doctoral dissertation) Retrieved from http://dareassociation.org/Terri.Robinett_Dissertation_2006.pdf

A study analyzed the similarities and differences in moral and reasoning stages between liberal and conservative voters. This study utilized the model of hierarchical complexity to relate various individual performances on multiple measures of moral reasoning. Although the model of hierarchical complexity did not predict political affiliation, it did support the notion that the test items were measuring moral reasoning stages, which provides support for Kohlberg's stage theory. Education-level and household income were found to be highly correlated and significant predictors of

political affiliation while level of religiosity was correlated with and found to be a significant predictor of one's identification as a liberal or a conservative. This study shows that stage and value (conservative versus liberal) are not always related.

17. Commons, M. L., & Miller, P. M. (2007). How early negative caregiving experiences relate to stage of attachment. *Behavioral Development Bulletin*, *13*, 1417. Retrieved from http://dareassociation.org/Papers/BDB%203.%20Early%20Negative%20Caregiving%20Experiences.pdf

The paper explained how negative behaviors have an effect on stages of adult development of attachment by using the model of hierarchical complexity. Extreme negative early experiences, such as abuse or neglect, often lead to arrested development in the subdomains, such as the interpersonal domain, in which such experiences occurred. The model of hierarchical complexity helped to explain how different stage behavior attachment issues should be responded to with different approaches. This showed how reinforcement contingencies (value) affect emotional stage.

18. Miller, P. M., & Commons, M. L. (2007a). Stages of infant development, as illustrated by responses to the peekaboo game in infants. *Behavioral Development Bulletin*, *13*, 1823. Retrieved from http://dareassociation.org/Papers/BDB%204.%20Infant%20stages%20peek-a-boo.pdf

In this paper, infant developmental changes were characterized by the model of hierarchical complexity. A possible developmental stage sequence for peek-a-boo and its comparison to the same stage sequences of other infant behaviors was discussed. This showed that the early stages in the model of hierarchical complexity can analyze and explain infant developmental behaviors.

19. Miller, P. M., & Commons, M. L. (2007b). How are the processes by which people become attached influenced by stage of development? *Behavioral Development Bulletin*, *13*, 2429. Retrieved from http://www.dareassociation.org/Papers/BDB%205.%20Processes_People%20Become%20Attached.pdf

Attachment processes are those events that are involved in developing and strengthening the attachment behaviors at different stages. The model of hierarchical complexity provided a rationale for specifying the order of development of attachment processes. It was used to examine the attachment process for children from age eight to ten years.

20. Commons, M. L., Goodheart, E. A., Pekker, A., Dawson, T. L., Draney, K., & Adams, K. M. (2008). Using Rasch scaled stage scores to validate orders of hierarchical complexity of balance beam task sequences. *Journal of Applied Measurement*, *9* (2), 182–199. Retrieved from http://dareassociation.org/Papers/Commons%20URM.pdf

The studies listed in the paper examine the relationship between the analytic basis underlying the hierarchies produced by the model of hierarchical complexity and the probabilistic Rasch scales that placed both participants and problems along a single hierarchically ordered dimension.

21. Day, J. M. (2010). Religion, spirituality, and positive psychology in adulthood: A developmental view. Journal of Adult Development, 17, 215–229.

The article considered contributions of religious commitment and spiritual practice to well-being and cognitive-developmental theoretical models. It also related bodies of empirical and clinical research with religious and spiritual development across the life cycle, with particular attention to questions related to positive adult development. The model of hierarchical complexity could be used to analyze religious related issue in respect of adult development.

22. Commons, M. L., Miller, P. M., Li, E. Y., & Gutheil, T. G. (2012). Forensic experts' perceptions of expert bias. *International Journal of Law and Psychiatry 35* (5–6), 362–371. Retrieved from http://dareassociation.org/Papers/CommonsMillerLiGutheil2012.pdf.

The purpose of this study was to find out how potentially biasing different situations were perceived to be. Another focus of the current study was to understand perceived bias from a psychophysical perspective. In the paper, the reason why some biases are easier to identify and overcome than others was identified. According the model of hierarchical complexity, more difficult items required higher stages. A factor analysis showed that the higher the stage, the less biasing they were perceived to be. This shows that the model of hierarchical complexity was useful in understanding the degree of bias in various situations.

23. Commons, M. L., Li, E. Y., Richardson, A. M., Gane-McCalla, R., Barker, C. D., & Tuladhar, C. T. (2014). Does the model of hierarchical complexity produce significant gaps between orders and are the orders equally spaced? *Journal of Applied Measurement*, *15* (4), 1–29.

The paper shows that, in the model of hierarchical complexity, there was no overlap between the Rasch-scaled item scores at one order of complexity, and those of the adjoining orders. There are "gaps" between the stages of performance on those items. And after testing for equal spacing between the orders of hierarchical complexity, it was found that the orders of hierarchical complexity were equally spaced.

24. Commons, M. L. & Tuladhar, C. T. (2014). Developmental behavior analytic therapy part I: easier done than said. *Behavioral Development Bulletin, 19* (2), 1–10.

Developmental behavior analytic therapy (DBAT) is the first behavioral therapy with developmental underpinnings. This paper introduces DBAT by presenting a composite case study. It also discusses the theoretical underpinnings of this therapy. DBAT aims to help individuals with behavioral problems change specific problem behaviors that consequently help them to lead more satisfying lives. It aims to alter specific behavioral problems,

because the biological susceptibility to such behavioral problems is a given. It is suggested that this therapy be used as an adjunct to conventional therapies that specialize in helping individuals cope with behavioral problems. DBAT is different from other contemporary or behavioral analytic therapies, as it integrates a behavioral developmental stage model, the model of hierarchical complexity (MHC), into its working. The foundation of this therapy is the theory that behavioral developmental stages and value of consequences of a behavior interact to predict an individual's behavior, and also suggests that behavioral problems affect both behavioral developmental stage and value of consequences.

25. Commons, M. L. & Tuladhar, C. T. (2014). Developmental behavior analytic therapy part ii: procedures and case studies, *Behavioral Development Bulletin, 19* (2), 11–21.

This paper discusses the case studies applying developmental behavioral analytic therapy (DBAT), a new behavioral therapy with developmental underpinnings. It also lays out the sequence of procedures of this therapy. The procedures have been illustrated with examples from six case studies of individuals who have undergone the therapy. It also presents the methodology and results of intervention using DBAT on those six individuals. With DBAT, five out of the six individuals achieved their target behaviors and increased their developmental stages. The positive results yielded from this small sample suggest potential benefit and success of DBAT therapy.

26. Commons, M. L., Commons-Miller, L. A. H., Salaita, R. J., & Tuladhar, C. T. (2014). Stars that crash. *Behavioral Development Bulletin, 19* (2), 100–110.

The present study introduces a model explaining what leads stars to crash and assesses risk factors that lead stars to crash in a sample of 18 celebrities who have had a downfall. Downfalls include alcoholism, drug abuse or addiction, mental illness, myriad relationship problems, death, suicide or other life-changing disasters. First, the paper theorizes that individuals' early environments and social forces, such as assortativeness and affiliation, contribute to their narcissistic traits. The model illustrates how these risk factors including narcissistic traits and the adult environments of stars lead them to engage in behaviors that lead to their downfalls. To examine the usefulness of this model, the paper examined the lives of famous celebrities (i.e., "stars") who had public downfalls ($n = 18$) using secondary sources. It assessed the risk factors involved in the crashing of stars. In concordance with the proposed model, results showed that what the majority of these cases had in common were: Atypical early environments, such as abandonment and trauma, over-indulgent or absent wealthy parents, or an early career; and adult environment conditions, such as colluding social groups and entourages. These factors could be linked to stars having extramarital affairs damaging their marriage or careers; bankruptcy; or alcohol and/or drug addiction. In some cases these factors have led to stars having accidents, or deaths. Furthermore, the study shows that there is a positive correlation between the number of risk factors present and the severity of the downfall of the stars.

27. Li, E. Y., Commons, M. L., Miller, J. G., Robinett, T. L., Marchand, H., Ost, C. M., & Ross, S. N. (2014). Relationship among measures within the social and moral development domain. *Behavioral Development Bulletin, 19* (2), 106–113.

This paper investigates using the model of hierarchical complexity (MHC) as a framework to study individual's stages of moral understanding. As an improvement from traditional stages of moral development, 15 stages of moral understanding were generated using the model of hierarchical complexity. Data were collected in four separate studies on how participants make choices in specific moral dilemmas. Each study presented five or six vignettes of arguments, each constructed to have different orders of hierarchical complexity. Participants rated the quality of arguments on a 1 to 6 scale. A Rasch analysis produced stage scores for each of the stories. The Rasch scores were regressed against the order of hierarchical complexity of each vignette. These were Counselor-Patient: $r(3) = .992$; Anti-Death-Penalty: $r(3) = .919$; Incest –No Report: $r(3) = .916$; Incest-Report: $r(3) = .624$. The result showed that Rasch scores of vignettes were predicted by their orders of hierarchical complexity, suggesting that the model of hierarchical complexity was a good framework to study stage of moral understanding.

28. Ross, S. N., Commons, M. L., Li, E. Y., Stålne, K., & Barker, C. D. (2014). Toward defining order 16 and describing its performance for the model of hierarchical complexity. *Behavioral Development Bulletin, 19* (3), 33–36.

We trace the first four years of the new theoretical discourse on the definition order 16 of hierarchical complexity. Tasks performed at this order are similarly classified as stage 16 performances. Until this current discourse began, the highest order identified using the MHC was order 15, named cross-paradigmatic. In different groupings, several MHC theorists have discussed the properties and definition of this new order. To this point, an explicitly collaborative effort has yet to be undertaken. To reach agreement on definition and properties of order 16 and task performances at that order will likely require us to agree on more complex than usual hierarchical complexity-based scoring criteria and inter-rater standards. To meet these new challenges, these criteria and standards must be precise enough, complex enough, and general enough to apply across the uncommonly disparate and high-level examples proposed thus far as performances at stage 16. Since these methodological foundations have not yet been developed, to date our discourse is comprised of some who consider the process of defining the new order and empirically demonstrating it further along than others do. This theoretical development terrain promise intense and promising work ahead on this breakthrough in applying the MHC, its contributions to behavioral development theory, and the measurement of the most complex human accomplishments recognized thus far.

29. Tuladhar, C. T., & Commons, M. L. (2014). Correspondence between some life-span stage theory developmental sequences of stages and levels. *Behavioral Development Bulletin. 19* (3), 24–27.

Good comparisons of development sequences have been made in the past. The model of hierarchical complexity is one developmental sequence which has often been compared to other developmental sequences including: Piaget & Inhelder (1969); Fischer & Bidell (1998); Colby and Kohlberg's (1987A, 1987B) 9 point stages and moral maturity scores (mms) of moral judgment. However, Colby and Kohlberg's 13 point scale has never been assessed in making comparisons to other scales. The current paper constructed a comparison table of all five models, including Colby and Kohlberg's 13 point scale, which together cover the developmental stages of an entire life-span. Adjustments had to be made to the 9 point and 13 point scales. The formula, OHC = 3 + 2 * (Stage of Colby & Kohlberg's), was introduced to demonstrate the relationship between the orders of hierarchical complexity and Kohlberg's stages of development.

30. Commons, M. L., & Jiang, T. R. (2014). Introducing a new stage for the model of hierarchical complexity: a new stage for reflex conditioning. *Behavior al Development Bulletin*, *19* (3), 1–8. Retrieved from http://baojournal.com/BDB%20WEBSITE/archive/BDB%2019.3-A01.pdf

The paper introduced the new stage 1 in the model of hierarchical complexity. The original sensory or motor stage 1was split into two stages. The new automatic stage 1 had respondent conditioning removed from it. By introducing the new stage, single cells behavior could be better understood. The revised sensory or motor stage 2 explained respondent conditioning (originally order 1) in non-human animals.

31. Commons, M. L., Miller, L. S., & Giri, S. (2014). A model of stage change explains the average rate of stage of development and its relationship to the predicted average stage ("smarts"). *Behavioral Development Bulletin*, *19* (3).

A number of different previous methods for measuring "smarts" have led to the model of hierarchical complexity (MHC), a context free neo-Piagetian mathematical model of behavioral complexity. It provides a way to classify tasks as to their hierarchical complexity. Using the model of hierarchical complexity, this study examines how differences in rate of stage change results in a difference in the highest average stage (smarts") attained by 70 year old adults. The average stage of development ("smarts") was shown to be predicted by the log of age with an *r* = .79. It uses data from Colby, Kohlberg, Gibbs, Lieberman (1983) to test the model. It also predicts that on the average there is one stage of development during adulthood.

32. Harrigan, W. J., & Commons, M. L. (2014). The stage of development of a species predicts the number of neurons. *Behavioral Development Bulletin*, *19* (3).

This paper asked the question: does energy density as a represented by number of neurons of organisms increase as stage of animals' increases? This paper also addresses the question of why brains evolved in the first place. The stage of development of several species was assessed. The model of hierarchical complexity successfully predicts the number of neurons in different species.

33. Giri, S. P, & Commons, M. L., Harrigan, W. J. (2014). There is only one stage domain, put in the published information. *Behavioral Development Bulletin*, *19* (3).

The study used the model of hierarchical complexity (MHC) to test the theory that different skills in development would develop in synchrony, allowing an individual to solve tasks from various domains using the same mental structure for each task. The MHC instruments used were the empathy, helper person, counselor patient, breakup, caregiver, algebra, balance beam, infinity and laundry instruments. The analysis showed that person scores on each instrument loaded on the first factor (92.156% of the variance). This agrees with the model of hierarchical complexity's assumption that there is only domain in development.

34. Commons, M. L., Giri, S. P. & Harrigan, W. J. (2014). The small effects of non-hierarchical complexity variables on performance. *Behavioral Development Bulletin*, *19* (3).

This study was designed to test the effect of small variables on task performance. The variables tested were hierarchical complexity, place in order, the number of calculations needs, the size of the numbers, and the causal variable position. Participants were asked to solve problems from task sequences from the Logic/mathematics/physical science subdomains. The results showed that order of hierarchical complexity has a very strong predictive role and accounts for most of the variance. The other variables only made very small contributions.

35. Commons, M. L., & Thexton, C. S. (2014). Stages of investing using the model of hierarchical complexity. *Behavioral Development Bulletin*, *19* (4).

Most theories and studies of decision making are a-developmental. However, there is ample evidence that there are differences in behavior on many decision-making tasks between children and adults. This paper asserts that within adults there are differences in behavior on many decision-making tasks and discusses investment as a decision-making task where differences in adult behavior can be analyzed. It presents an argument that stage theory can determine investment behavior. The major properties of investment behavior are *a)* how many variables a person can look at and *b)* whether a person can compare systems and understand that regulations are incomplete and not consistent. We propose that the rational theories of investing fail because most economic theories assume perfectly rational players in the market place. One of the major reasons that private investors do terribly in managing and investing money is the inadequate stage development of the investors on the task of investing.

36. Alves, J. (2014). Depression, the loss of value and its relationship to stage. *Behavioral Development Bulletin*, *19* (4).

Depression is almost always accompanied by the loss of value for previously salient reinforcement. But do the sources of reinforcement value change with stage? At every stage, punishment

produces depression is some fashion. Punishment has two effects. First is during acquisition, punishment strengthens competing behaviors by being paired with the elicitor of the operant behavior that interferes with the punished behavior. Second during allocation it changes the value of the punished behavior relative to the other competing behaviors. For example, primary stage 8 performing people think only about themselves. At that primary stage, people receive a lot of punishment because they do not take other people's perspectives and integrate those perspectives with their own. They are the second lowest stage criminals. Non-psychotic rapists are an example. They are not severely depressed but moderately depressed. They are nasty and there is a tendency toward aggression. If you ask them, they say that the world sucks. They see themselves as victims.

37. Commons-Miller, L. A. H., Commons, M. L., Li, E. Y., Miller, P. M., Golino, H. F. & Tuladhar, C. T. (in press). Stage of pricing strategy predicts earnings: a study of informal economics.

The study examined the relationship between stages of development on economic tasks and income of the people being studied. It shows that the order of hierarchical complexity was a better predictor of income and indicates that social stratification will continue to persist as differences in individual's developmental stages will persist.

38. Harrigan, W. J. (in press), November. Replacing Maslow's needs hierarchy with an account based on stage and value.

An all-encompassing model of value and stage was applied to interpret Maslow's "needs" hierarchy model. The model of hierarchical complexity was used to expand the understanding of Maslow's model. It was suggested that Maslow's hierarchy of needs be interpreted from the perspective of stage and value.

39. Commons, M. L., Grotzer, T. A. & Davidson, M. N. (in press). The sufficiency of reinforcing problem solutions for transition to formal operations.

A racially and socio-economically integrated population of fifth and sixth grade students was repeatedly presented with formal-operational causality problems. With problem presentation alone and with problem presentation and feedback, no significant change occurred. The transition to the formal stage performance accelerated significantly only when correct answers were reinforced. There were 75% of the participants reaching the formal stage on the problem they were trained. This finding supports the sufficiency of reinforcement for producing formal stage performance.

40. Commons-Miller, L. A. H, Commons, M. L, Miller, P. M., Golino, H. F., & Li, E. Y. (in Press). Stage of pricing strategy predicts earnings: a study of informal economics.

Social stratification is a significant moral issue mostly driven by high levels of income disparity. A common notion is that such income disparity can be reduced by creating equal opportunity of education for all individuals. However, this study shows that developmental stage is a better predictor of income and indicates that social stratification will continue to persist as differences in individuals' developmental stages will persist. This cross-cultural study examines the relationship between stages of development on economic tasks and income of the people being studied. Two groups of people were studied: people who sell things on the sidewalks (peddlers) and people who transport goods (carters). Participants were from Brazil and the United States. Studying informal economies across cultures allowed us to test the stage of pricing strategies used by people of varying education levels and at different behavioral stages of development ranging from primary to metasystematic. The purpose of this study was to determine the extent to which the behavioral developmental stage of economic reasoning affected income obtained. Three quasi-independent variables were examined: the behavioral stage of the person's economic behavior, the country in which they lived, and how much schooling they had. It was found that the developmental stage of participants' pricing strategies, correlated most with how much they earned. The developmental stage was a better predictor of income than education.

41. Day. J, (in press). On cognitive complexity, stage, and value, in religious cognition.

This study examined whether level of cognitive complexity in religious cognition, using measures rooted in the model of hierarchical complexity, mediates the relationship between level of general cognitive development, and religious judgment, in children and adolescents. The study was conducted with 189 children and adolescents drawn from Catholic schools in France. General cognitive development level was measured using the wisc and wais. Level of cognitive complexity in religious cognition was measured using the *Religious Cognition Questionnaire: Pastor-Parishioner Scenario* (rcq). Religious judgment levels were measured using the *Religious Cognition Questionnaire: Pastor-Parishioner Scenario* (rcq). Results indicate that the relationship between iq and religious judgment is mediated by level of complexity in religious cognition. The results provide further empirical evidence for the conceptual validity and research utility of the model of hierarchical complexity in the domain of religious cognition, and the power of the concept of complexity in religious cognition for explaining relationships between general cognitive development using iq measures and other domains where the judgment of social, moral, and philosophical issues are concerned. The results also provide further empirical evidence for the distinctive, and relatively advanced, capacity of "gifted" young people to think critically about religious and philosophical issues.

42. Görtz, D. P., & Commons, M. L. (in press). The stage-value model: implications for the changing standards of care.

The standard of care is a legal and professional notion against which doctors and other medical personnel are held liable. The standard of care changes as new scientific findings and technological innovations within medicine, pharmacology, nursing

and public health are developed and adopted. This study consists of four parts. Part 1 describes the problem and gives concrete examples of its occurrence. The second part discusses the application of the model of hierarchical complexity on the field, giving examples of how standards of care are understood at different behavioral developmental stage. It presents the solution to the problem of standards of care at a paradigmatic stage 14. The solution at this stage is a deliberative, communicative process based around why certain norms should or should not apply in each specific case, by the use of so-called meta-norms. Part 3 proposes a cross-paradigmatic stage 15 view of how the problem of changing standards of care can be solved. The proposed solution is to found the legal procedure in

Table 5. the correlation coefficients of the model of hierarchical complexity and each instrument's Rasch data

study names	measures	r for predicting Rasch difficulty	factor loading	
			1	2
there is only one stage domain	helper person	$r(40) = 0.977$	0.969	0.087
there is only one stage domain	laundry	$r(111) = 0.964$	0.999	0.028
there is only one stage domain	empathy	$r(22) = 0.910$	0.939	−0.321
there is only one stage domain	counselor patient	$r(30) = 0.934$	0.952	−0.158
there is only one stage domain	breakup dilemma	$r(21) = 0.835$	0.954	−0.239
there is only one stage domain	caregiver	$r(42) = 0.711$	0.964	−0.107
there is only one stage domain	algebra	$r(40) = 0.966$	0.975	0.143
there is only one stage domain	balance beam	$r(51) = 0.980$	0.919	0.433
there is only one stage domain	infinity [reference]	$r(52) = 0.912$	0.966	0.137
developmental task action sequence in the development of communication	autism developmental task sequence (ADTS)	$r(42) = 0.893$	—	
why there are upper limits to the stage of development on one's best task and best domain	$\log_2$(age) predicted moral maturity score	$r(225) = 0.764$ $p = 0.000$	—	
the stage of development of a species predicts the number of neurons	stage of development predicts number of neurons	linear: $R^2 = 0.580$ power function: $R^2 = 0.764$	—	
forensic experts' perceptions of expert bias	the Rasch scores of the items in the instrument and the orders of hierarchical complexity of the items	$r(16)=0.698$	—	

each case on well-established behavioral laws. We maintain that such a behavioristic, scientifically based justice would be much more proficient at effecting restorative legal interventions that create desired behaviors. The forth part address a chaos theory of why parts one, two and three will occur. Simply stated, it sets forth that the change is driven by attractors that produce short and long term reinforcing benefits. These attractors select the behaviors that obtain them even in a noisy environment where is are ups and down in policy.

43. Görtz, D. P. & Commons, M. L. (in press). The societal attractors towards flatter organization in the 21st century in medicine, law, business, public administration and politics.

Organizations in general are likely to become flatter in the 21st century, which also affects the future of institutions of medical care, psychiatric care and forensics. This trend can be seen already today. This paper suggests eight long-term "attractors" (rather than linear short-terms trends) that lead to flatter organizations. Organizations and the need for management and bureaucracy are discussed from an informational perspective. It is argued that management and hierarchical bureaucracy are necessary because information about specific behaviors needs to be processed by organizations, but that informational efficiency simultaneously implies the striving towards the lowest possible number of layers. Then the eight attractors (long-term trends) are presented: *a)* the cultural evolution of information management; *b)* scientific support for new forms of management; *c)* social media technology; *d)* social innovation of new management forms; *e)* new production and distribution chains; *f)* the strong growth of the world market; *g)* the radically disruptive technological development; *h)* the re-in-

tegration of business, politics and civil sphere. Some consequences in medicine, psychiatry and forensics are presented, including the new forms of treatment in psychiatry where psychiatric treatment is likely to become increasingly merged with the real everyday life situations of the patients.

» TASK DIFFICULTY PREDICTION

The model of hierarchical complexity (MHC) measures stage of development within and across domains, allowing comparison between domains. The table below illustrates the r scores for predicting Rasch difficulty and it corresponding factor loadings by using the model of hierarchical complexity as scale. The data shows that the model of hierarchical complexity successfully predicted task difficulties.

» DISCUSSION

The list of advances shows how the model of hierarchical complexity has been developing. It also shows how the theory has been applied to numerous topics. The model of hierarchical complexity is now being used in all inhabited continents of the world except Africa. Because the model is so simple and is based on analysis of tasks and not just performances, it is dynamic. This has been shown by the addition of the new orders of hierarchical complexity and the revision of scoring of examples to illustrate other orders. Where might the model of hierarchical complexity go? There seem to be two branches of investigation. One is the construction of instruments in an ever widening set of domains. Because much of the work so far has been done in adulthood, the 'new frontier' has been to conduct similar kinds of studies with children and adolescents and also with non-human animals. Another is the scoring of interviews, speeches and text, also in an expanding variety of domains.

The application of the model of hierarchical complexity has added, along with other stage approaches, to the understanding of the interaction of value and stage; to organizational development, education, and the development of science and culture. Each of these areas will be briefly discussed next.

Organizational development

For understanding organizational development, the model of hierarchical complexity is used to analyze the structure of organizations and to match employees' stage of performance with the stages required to perform the tasks demands in their job descriptions. The degree of the match is used to select such employees for organizations. This work is ongoing. In the newest work, stage of performance is being combined with six factor loadings of interest.in the Holland scale (Holland, 1973) Interest constitutes the "value" part of the consequences of task engagement and completion. When using interest's factor loadings combined with stage of performance, a match between the position demands in and organization is enhanced greatly over just using personality matching. The squared differences between each variable are minimized by choosing the right person. The process can also be used to find the right date for people seeking relationships

Education

One of the exciting applications has been to the area of computer run education. An example would be the use of computer run education to teach the model of hierarchical complexity. Another use has been to understand how to arrange items in a computer run curriculum so that they have a higher chance of being ordered properly by increasing difficulty. Using adaptive means, and using age as a predictor of stage of performance on an item, one would have the program start at the appropriate age, as a first approximation. If the person preformed correctly, the next item in the developmental task sequence would be presented. If they did not get the item predicted by age, a less difficult item in the sequence would be chosen. The adaptive way would maximize the rate of acquisition of material from the next stage.

Terrorisms and politics

Currently, as Commons (Commons & Goodheart, 2007) has argued, government-building in troubled areas has simply proceeded by one country (often the United States) attempting to impose the current system of government as is in the United States on another country. Commons has argued that the types of governmental structures that countries have are stage-related, and that to impose a system that is not a good match on a country will inevitably fail. Instead, government-building would have the tasks match the same stage at which enough of the powerful part of the society is operating at, or perhaps just slightly above that stage. The unrealistic expectations of helper governments would be replaced by successful interventions at each progressive stage.

Table 6. societal developmental periods and their correspondence to developmental stages

societal development periods	corresponding stages
ultra-traditionalism	transition to concrete
traditionalism	concrete
transition to modernism	abstract
early modernism	formal
modernism	systematic
transition to ultra-modernism	metasystematic
ultra modernism	paradigmatic
advanced ultra-modernism	cross paradigmatic

Development of science

There are a number of applications to the development of science. One of the applications to science that proceeds rather directly out of the model of hierarchical complexity is the proposal for stacked neural networks. One of reasons to propose such stacked neural networks is that existing neural network models are not articulated enough to take on complex learning. The only ones that exist now have two layers, not counting the internal layers within a network. Networks that are more multilayered can address issues at each developmental stage and should therefore be able to better model intelligence of a variety of animals and of humans. For example, in a nominal order 5 task, a four layered stack could *understand* single words. It could respond to the *meaning* of such words.

Stacked neural networks use the model of hierarchical complexity (Commons & Chen, 2008) to accomplish the following tasks: model human development and learning; reproduce the rich repertoire of behaviors exhibited by humans; allow computers to mimic higher order human cognitive processes and make sophisticated distinctions between stimuli; and allow computers to solve more complex problems.

Development of culture

In very condensed form, the societal developmental periods have been analyzed as to their corresponding stage of development. Remember stage change can only take place one stage at a time. There cannot be any skipping of stages because the higher stage always requires proficient next lower stage behavior. The suggested correspondence between type of societal development, and stage are shown in Table 5.

Above, we outline some of the current work using the model of hierarchical complexity. Because the model of hierarchical complexity is an open system, there is no way to predict what the future applications will be.   ■

REFERENCES

Alves, J. (2014). The loss of value in depression and stage. *Behavior Development Bulletin*, *19*(4).

Asimov, I. (1951). Foundation, Gnome Press.

Biggs, J., & Collis, K. (1982). A system for evaluating learning outcomes *The SOLO taxonomy*. New York: Academic Press.

Commons (1976a). The structural analytic stage. A talk given to the Mathematics Department, Northern Michigan University, Marquette, Michigan.

Commons, M. L., & Chen, S. J. (2014, March). *Stacked neural network.* Poster presented at the 2014 Annual Meeting of the Eastern Psychological Association, Boston, MA.

Commons, M. L. & Goodheart, E. A. (1999). The philosophical legacy of behaviorism. In B.A. Thyer (Ed.). *The philosophical foundations of behaviorism* (pp. 9–49). England: Kluwer Academic. Publishers.

Commons, M. L. & Richards, F. A. (2002). Organizing components into combinations: how stage transition works. *Journal of Adult Development*, *9*(3), 159–177.

Commons, M. L. & Richards. F. A. (2003). Four post formal stages handbook of adult development. *The Springer Series in Adult Development and Aging*, 2003, 199–219. Retrieved from http://www.tiac.net/~commons/Four%20Postformal%20Stages.html

Commons, M. L. & Rodriguez, J. A. (1993). The development of hierarchically complex equivalence classes. *Psychological Record 43*, 667–697.

Commons, M. L. & Rodriguez. J. A. (1990). "Equal access" without "establishing" religion: the necessity for assessing social perspective-taking skills and institutional atmosphere. *Developmental Review*, *10*, 323–340. Retrieved from http://dareassociation.org/Papers/Equal%20Access%20without%20Establishing%20religion1990.pdf

Commons, M. L., & Tuladhar, C. T. (2014). Developmental behavior analytic therapy part ii: procedures and case studies. *Behavioral Development Bulletin*, *(19)*2, 11–21.

Commons, M. L., & Tuladhar, C. T. (2014). Developmental behavior analytic therapy i: easier done than said. *Behavioral Development Bulletin, (19)*2, 1–10.

Commons, M. L., & White, M. S. (2003). A complete theory of tests for a theory of mind must consider hierarchical complexity and stage. *Behavioral and Brain Sciences, 26*(5), 606–607.

Commons, M. L., & Wolfsont, C. A. (2002). A complete theory of empathy must consider stage changes. *Behavioral and Brain Sciences*, *25*(1), 30–31.

Commons, M. L. (1991). A comparison and synthesis of Kohlberg's cognitivedevelopmental and Gewirtz's learningdevelopmental attachment theories. *J. L. Gewirtz & W. M. Kurtines (Eds.) Intersections with attachment.* 257291.

Commons, M. L. Commons-Miller, L.A.H., Salaita, R.J., & Tuladhar, C.T. (2014). Stars that crash. *Behavioral Development Bulletin, 19*(2), 100–110.

Commons, M. L., & Goodheart, E. A. (2007). Consider stages of development in preventing terrorism: Does government building fail and terrorism result when developmental stages of governance are skipped? Journal of Adult Development, 14, 91–111.

Commons, M. L., & Jiang, T. R. (2014). Introducing a new stage for the model of hierarchical complexity: a new stage for reflex conditioning. *Behavior Development Bulletin*, *19*(3), 1–8. Retrieved from http://baojournal.com/BDB%20WEBSITE/archive/BDB%2019.3-A01.pdf

Commons, M. L., & Miller, L. S. (2014. May). *Upper limits to the stage of development.* Posters presented at the annual meeting of the Society for Research in Adult Development, Salem, MA.

Commons, M. L., & Miller, P. M. (2007). How early negative caregiving experiences relate to stage of attachment. *Behavioral Development Bulletin*, *13*, 1417. Retrieved from http://dareassociation.org/Papers/BDB%203.%20Early%20Negative%20Caregiving%20Experiences.pdf

Commons, M. L., & Richards, F. A. (1978). The structural analytic state of development: A Piagetian post-formal operational stage. Paper presented at the meeting of the *Western Psychological Association*, San Francisco, April 1978.

Commons, M. L., & Thexton, C. S. (2014). Stages of investing using the model of hierarchical complexity. *Behavior Development Bulletin*, *19*(4).

Commons, M. L., Giri, S. P., & Harrigan, W.J. (2014). The small effects of non-hierarchical complexity variables on performance. *Behavior Development Bulletin*, *19*(3).

Commons, M. L., Goodheart, E. A., Pekker, A., Dawson, T. L., Draney, K., & Adams, K. M. (2008). Using Rasch scaled stage scores to validate orders of hierarchical complexity of balance beam task sequences. *Journal of Applied Measurement*, *9*(2), 182–199. Retrieved from http://dareassociation.org/Papers/Commons%20URM.pdf

Commons, M. L., Grotzer, T. A., & Davidson, M. N (in press). The sufficiency of reinforcing problem solutions for transition to formal operations. .

Commons, M. L., Li, E.Y., Richardson, A.M., Gane-McCalla, R., Barker, C.D., & Tuladhar, C.T. (2014). Does the model of hierarchical complexity produce significant gaps between orders and are the orders equally spaced? *Journal of Applied Measurement*, *15*(4), 1–29.

Commons, M. L., Miller, L.S., & Giri, S. (2014). A model of stage change explains the average rate of stage of development and its relationship to the predicted average stage ("smarts"). *Behavior Development Bulletin*, *19*(3).

Commons, M. L., Miller, P. M., & Kuhn, D. (1982). The relation between formal operational reasoning and academic course selection and performance among college freshmen and sophomores. *Journal of applied developmental psychology*, *3*, 1–10. Retrieved from http://www.dareassociation.org/Papers/The%20Relation%20between%20formal%20operational%20reasoning1982.pdf

Commons, M. L., Miller, P. M., Li, E. Y., & Gutheil, T. G. (2012). Forensic experts' perceptions of expert bias. *International Journal of Law and Psychiatry 35*(5–6), 362–371.

Commons, M. L., Richards, F.A., & Kuhn, D. (1982). Systematic and metasystematic reasoning: a case for levels of reasoning beyond Piaget's stage of formal operations. *Child Development, 3*, 1058–1069 Retrieved from http://dareassociation.org/Papers/Systematic%20and%20Metasystematic%20reasoning1982.pdf

Commons, M.L., Rodriguez, J.A., Adams, K.M., Goodheart, E.A., Gutheil, T.G., & Cyr, E.D. (2006). Informed consent: do you know it when you see it? evaluating the adequacy of patient consent and the value of a lawsuit. Psychiatric Annals, 36, 430–435. Retrieved from http://dareassociation.org/Papers/Informed%20Consent%202006.pdf

Commons, M. L., Miller, P. M., & Kuhn, D. (1982). The relation between formal operational reasoning and academic course selection and performance among college freshmen and sophomores. *Journal of Applied Developmental Psychology, 3*, 1–10.

Commons, M. L., Krause, S. R., Fayer, G. A., & Meaney, M. (1993). Atmosphere and stage development in the workplace. In J. Demick & P. M. Miller (Eds.). *Development in the workplace*, 199–220. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. Retrieved from http://dareassociation.org/Papers/CommonsKrauseFayerMeaney1993.pdf

Commons. M. L., Lee. P., Gutheil, T. G., Rubin, E., Goldman, M., & Appelbaum, P. L. (1995). Moral stage of reasoning and the misperceived "duty" to report past crimes. *International Journal of Law and Psychiatry, 18*(4), 415–424. Retrieved from http://dareassociation.org/Papers/Moral%20State%20of%20Reasoning1995.pdf

Commons-Miller, L. A. H, Commons, M. L, Mille, M. P., Golino, H. F., & Li, E. Y. (in press). Stage of pricing strategy predicts earnings: a study of informal economics.

Day, J. M. (2010). Religion, spirituality, and positive psychology in adulthood: a developmental view. Journal of Adult Development, 17, 215–229.

Day. J, (in press). On cognitive complexity, stage, and value, in religious cognition.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review 87*: 477–531.

Görtz, D. P. & Commons, M. L. (in press). The societal attractors towards flatter organization in the 21st century in medicine, law, business, public administration and politics.

Görtz, D. P. Commons, M. L. (in press). The stage-value model: implications for the changing standards of care.

Harrigan, W. J. (in press). Replacing Maslow's needs hierarchy with an account based on stage and value.

Harrigan, W. J., & Commons, M. L. (2014). The stage of development of a species predicts the number of neurons. *Behavior Development Bulletin*, *19*(3).

Harrigan, W. J., Giri, S. P., & Commons, M. L. (2014. May). *There is only one stage domain*, Posters presented at the annual meeting of the Society for Research in Adult Development, Salem, MA.

Heijenoort, Jean van (1967). *A Source Book in Mathematical Logic, 1879–1931.* Harvard University Press.

1930. "Some metamathematical results on completeness and consistency," 595–96. Abstract to (1931).

1930. "The completeness of the axioms of the functional calculus of logic," 582–91.

1931. "On formally undecidable propositions of *Principia Mathematica* and related systems," 596–616.

1931a. "On completeness and consistency," 616–17.

Holland, J. L. (1973). Making vocational choices: a theory of careers. *Englewood Cliffs: Prentice-Hall*.

Inhelder, B., & Piaget, J. (1958). The growth of logical thinking from childhood to adolescence: an essay on the development of formal operational structures. *A. Parsons & S. Seagrim, Trans.* New York, NY: Basic Books (original work published 1955)

Kuhn, D., & Brannock, J. (1977). Development of the isolation of variables scheme in experimental and 'natural experiment' contexts. *Developmental Psychology*, *13*(1), 9–14.

Kuhn, D., Ho, V., & Adams, C. (1979). Formal reasoning among pre- and late adolescents. *Child Development*, *50*(4), 1128–1135.

LaLlave, J. (2005). The acceptability of arguments in favor of and against the Iraq War. *Conflict & Communication Online*, *4*(2), 1–16. Retrieved from http://dareassociation.org/Papers/Iraq%20War.LaLlave.pdf

Li, E. Y., Commons, M. L., Miller, J. G., Robinett, T. L., Marchand, H., Ost, C. M., & Ross, S. N. (2014). Relationship among measures within the social and moral development domain. *Behavioral Development Bulletin*, *19*(2), 106–113.

Li, E. Y., Commons, M. L., Miller, J. G., Robinett, T. L., Marchand, H., Ost, C. M., & Ross, S. N. (2014). Relationship among measures within the social and moral development domain. *Behavioral Development Bulletin*, *19*(2), 106–113.

Miller, P. M., & Commons, M. L. (2007a). Stages of infant development, as illustrated by responses to the peekaboo game in infants. *Behavioral Development Bulletin, 13*, 1823. Retrieved from http://dareassociation.org/Papers/BDB%204.%20Infant%20stages%20peek-a-boo.pdf

Miller, P. M., & Commons, M. L. (2007b). How are the processes by which people become attached influenced by stage of development? *Behavioral Development Bulletin, 13*, 2429. Retrieved from http://www.dareassociation.org/Papers/BDB%205.%20Processes_People%20Become%20Attached.pdf

Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica, 32*, 301–345.

Retrieved from http://dareassociation.org/Papers/CommonsMillerLiGutheil2012.pdf

Robinett, T. (2006). Moral reasoning and political affiliation in liberal and conservative voters: applying a model of hierarchical complexity (Doctoral dissertation). Retrieved from http://dareassociation.org/Terri.Robinett_Dissertation_2006.pdf

Rodriguez, J. A. (1992). *The adult stages of social perspective taking: Assessment with the doctor-patient problem*. Unpublished doctoral dissertation. Harvard Graduate School of Education.

Ross, S. N., Commons, M. L. , Li, E. Y., Stålne, K., & Barker, C. D. (2014). Toward defining order 16 and describing its performance for the model of hierarchical complexity. *Behavioral Development Bulletin, 19*(3), 33–36.

Ross. S. (2006). Effects of a structured public issues discourse method on the complexity of citizens' reasoning and local political development (Doctoral dissertation). Retrieved from http://dareassociation.org/Papers/Sara%20Nora%20Ross%20Dissertation.doc

Sonnert, G., & Commons, M. L. (1994). Society and the highest stages of moral development. *Politics and the Individual*, *4*(1), 31–55. Retrieved from http://dareassociation.org/Papers/Society%20and%20Highest%20Moral%20Dev.html

Tuladhar, C. T., & Commons, M. L. (2014). Correspondence between some life-span stage theory developmental sequences of stages and levels. *Behavioral Development Bulletin*. *19*(3), 24–27.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology, 20*(3), 273–281.

There is only one stage domain

Sagun Giri[1], Michael Lamport Commons[2], and William Joseph Harrigan[3]

1 Dare Institute
2 Harvard Medical School
3 Harvard University

ABSTRACT

The study used the model of hierarchical complexity (MHC) to test the theory that different domains in development would develop in synchrony, allowing an individual to solve tasks from various domains using the same mental structure for each task. The MHC instruments used were the empathy, helper person, counselor patient, breakup, caregiver, algebra, balance beam, infinity and laundry instruments. The instruments can be categorized as belonging to two different subdomains, the social subdomain, and the logic/mathematics/physical sciences subdomain. Instruments in this social subdomain measure developmental stage in a variety of social contexts. These social contexts included empathy for person after an accident, guidance and assistance by a helper, counseling patients, understanding romantic breakups, and caring for children and infants. The other subdomain is composed of mathematical (algebra & infinity), logical (laundry), and physical science (balance beam). In order to conclude how related the performances were, three analyses were carried out.

First, Rasch analysis yielded person scores akin to person stage scores. Second, regression analysis was conducted to assess how well the order of hierarchical complexity (OHC) of the items predicted the Rasch difficulty of the items. Third a principal axis factoring was carried out with the person Rasch scores for every instrument. Irrespective of domains, if each instrument loaded on the first factor with all the factor scores over .7 and if the first factor accounted for more than 70% of the variance, then that would show that all instruments were part of a single domain. In each case the MHC accounted for a large amount of variance with r values over .7.

The principal axis factoring showed that person scores on each instrument loaded on the first factor (90.51% of the variance). All the factor scores on the first factor were over .85. There were very low loadings only on the second factor (4.947% of the variance). This implies that the instruments from the social subdomain and instruments from the logic/mathematics/physical sciences belong to a single domain.

KEYWORDS: synchrony, one domain, model of hierarchical complexity

GILLIGAN (1992) CLAIMED THAT moral judgment and care formed different domains. Perry's 1968 study, 'Forms of Intellectual and Ethical Development in the College Years' explored the possibility of two domains: epistemological and meta-ethical development. Erdynast, Armon and Nelson's 1978 study Cognitive-Developmental Conceptions of the True, the Good and the Beautiful researched across three domains: conceptions of factual reality, conceptions of the good, and conceptions of the beautiful. Erdynast and Chen's (2014) study on the 'Relations between Adult Developmental Conceptions of the Beautiful and Moral Development' claimed that there are four theoretically distinct domains. These four domains are: I) the real, II) the good, III) the just and right and IV) the beautiful.

Kegan (1979), Kohlberg, Levine & Hewer (1983) and Commons & Richards (2003) however, asserted that there was just one domain. This study used a factor analysis of nine very different stage instruments to see how many factors are found. If the factor scores of the nine instruments loaded highly on the first factor, i.e. the first factor accounted for greater than 70% of the variance, then that would mean that all instruments were part of a single domain. This would mean that task performances across subdomains form a single domain, supporting Inhelder and Piaget's (1958) theory of synchrony of development.

Inhelder & Piaget (1958) found synchrony in the logical, mathematical and physical sciences domains. For decades this theory has been studied. The findings have been largely in contrast with the theory of synchrony. Fischer (1980) points out that these studies not only showed that décalage, also known as asynchrony, was the norm, but that the amount of asynchrony was often great. These studies do not make clear how tight concurrence must be

to support synchrony. They did not construct instruments to do the studies to show synchrony between domains.

» REFLECTIVE ABSTRACTION, HIERARCHICAL INTEGRATION, AND HIERARCHICAL COMPLEXITY

Many cognitive-developmental researchers assert that development in different knowledge domains does not necessarily proceed at the same rate (Fischer & Bidell, 1998; Lourenco & Machado, 1996). However, there is still considerable disagreement about whether development in different domains can be characterized in terms of a single generalized process. Domain theorists argue that different processes apply in different knowledge domains (Kohlberg, 1969; Larivee, Normandeau, & Parent, 2000; Turiel, 1980). Others, though they acknowledge that unique structures and processes are associated with particular domains also argue that a single general developmental process applies across domains (Case, Okamoto, Henderson, & McKeough, 1993; Fisher & Bidell, 1998). Piaget (2000) called this general process of reflective (or reflecting abstraction) abstraction, through which the actions of one developmental level become the subject of the actions of the subsequent level. The product of reflective abstraction is *hierarchical integration*. In conceptual development, hierarchical integration is observable in the concepts constructed at a new level by coordinating (or integrating) the conceptual elements of the prior level. These new concepts are said to be more hierarchically complex than the concepts of the previous level, in that they integrate earlier knowledge into a new form of knowledge. For example, independent conceptions of *play* and *learning* constructed at one complexity level are integrated into a conception of *learning as play* at the next complexity level (Dawson-Tunik, 2004a). Though it builds on the original *play* and *learning* concepts, the new concept cannot be reduced to the original *play* and *learning* elements. Not only is there a new concept in the recognition that learning can be playful, but the individual meanings of the elements *learning* and *play* have changed. The meanings of *learning* and *play* have changed in that each now incorporates some of the meaning embedded in the new construction. The concept of *learning* now includes *play* as a component, and the concept of *play* includes *learning* as a component.

A number of researchers have described developmental sequences that elaborate the basic notion of hierarchical integration, including Fischer (1980), who has emphasized the development of skill hierarchies in particular contexts. Commons and his colleagues (Commons et al., 1998) have described a task structure hierarchy. Pascual-Leone and Goodman (1979) have focused on the growth of mental attention and memory capacity. Case (1991) has described the development of memory capacity ad associated processing structures. Finally Demetriou and Valanides (1998) have described hierarchical development in terms of processing functions.

Table 1. the 17 known orders of hierarchical complexity

order	
number	name
0	computational
1	automatic
2	sensory or motor
3	circular sensory motor
4	sensory-motor
5	nominal
6	sentential
7	preoperational
8	primary
9	concrete
10	abstract
11	formal
12	systematic
13	metasystematic
14	paradigmatic
15	crossparadigmatic
16	meta-crossparadigmatic

Not only are there definitional correspondences among analogous levels described by Commons, Fischer, and Piaget, there is empirical evidence of correspondences between complexity levels, skill levels, and orders of hierarchical complexity and at least three domain-based systems, including Kitchener and King's (Dawson, 2002b; King, Kitchener, Wood, & Davison, 1989; Kitchener & King, 1990; Kitchener, Lynch, Fischer, & Wood, 1993) stages of reflective judgment, Armon's good life stages (Dawson, 2002a), Perry's stages of epistemological development (Dawson, 2004), and Kohlberg's moral stages (Commons et al., 1989; Dawson & Gabrielian, 2003; Dawson, Xie, & Wilson, 2003). These correspondences suggest that, as a community, this group of developmental researchers is moving toward a consensus regarding the detection and aspects of the definition of developmental stages.

» THE MODEL OF HIERARCHICAL COMPLEXITY

The nine instruments used in the study were from logic/mathematics/physical sciences subdomain and the social (social/moral/caring/empathy/perspective-taking/informed-consent) subdomain. These instruments were based on the model of hierarchical complexity. The model of hierarchical complexity (MHC) is a non-mentalistic, neo-Piagetian mathematical model (Krantz, Luce, Suppes, & Tversky, 1971; Luce & Tukey, 1964). MHC allows for the measurement of stage performance. It deconstructs tasks into the actions that must be done at each order. This is to build the behavior needed to successfully complete a task.

MHC provides an analytic *a priori* measurement of the difficulty of task actions. The difficulty is represented by the orders of hierarchical complexity (OHC) (Commons & Pekker, 2008). There are 17 known orders of hierarchical complexity. This is shown in Table 1

Hierarchical complexity describes a form of information that is different from traditional information theory (Shannon & Weaver, 1948) in which information is coded as bits that increase quantitatively with the amount of information. Theorem 4 of the model (Commons, et al, 1998) shows that every task action has an order of hierarchical complexity associated with it. The ideal correct task actions may be classified as to their order of hierarchical complexity. The tasks actions may address every experimental task, every clinical test item that has a difficulty associated with it, every behavior, developmental task, survey item, and statement made by people regardless of the content or context. Each task action will have a difficulty of performance associated with it.

A task action is defined as more hierarchically complex when *1)* A higher-order task is defined in terms of two or more tasks at the next lower order of hierarchical complexity, *2)* Higher-order tasks organize the lower order actions and *3)* The lower order tasks are coordinated non-arbitrarily, not just put together as an arbitrary chain. This is illustrated schematically in Figure 1.

» THE RASCH MODEL

Whereas they are well-known in psychometric circles, Rasch's (1980), models for measurement have been employed by developmental psychologists only recently (Andrich & Constable, 1984; Bond, 1994; Dawson, 1998, 2000, 2002; Draney, 1996; Muller, Sokol, & Overton, 1999; Wilson, 1984). These models are designed specifically to examine hierarchies of person and item performance, displaying both person proficiency and item difficulty estimates along a single interval scale (logit scale) under a probabilistic function. In addition, they can be employed to test the extent to which items or scores conform to a theoretically specified hierarchical sequence. A central tenet of stage theory is that cognitive abilities develop in a specified sequence, making the statistical tests implemented in a Rasch analysis especially relevant to understanding stage data. The Rasch model permits researchers to address questions like, "Are all single abstractions items more difficult than all representational systems level and less difficult than all abstract mappings items?" Moreover, the detailed information about item functioning and individual performances provided by the software makes it possible to simultaneously examine group and individual effects. These properties make Rasch models uniquely suitable for the investigation of many developmental phenomena.

The Rasch (1980) model uses logistic regression to minimize person and item error simultaneously. The model allows researchers to convert raw scores into equal interval linear scales. The item scores on the right represent how difficult the item was. The person scores on the left represent how good a person was at dealing with the item difficulty. The model also produces an objective, additive, and one dimensional scale. These are some of its advantages over other scaling techniques in measuring stage.

It is beyond the scope of this paper to provide a comprehensive account of the Rasch model, though we do attempt to provide enough information to allow readers who are unfamiliar with the model to follow the results of the analysis. For an introduction to the Rasch model, see Bond and Fox (2001), Rasch (1980), Smith (2004), and Wilson (2005).

A higher order action is:

1) defined in terms of the task actions from the **next lower order** of hierarchical complexity.

2) The higher order task action organizes two or more next lower order of hierarchical complexity.

3) The ordering of the lower task actions have to be carried out **non-arbitrarily**.

Figure 1. the three axioms of the MHC

physical sciences subdomain. The instruments were categorized as either belonging to the **mathematical** (Algebra & Infinity), **logical** (Laundry version of the Inhelder and Piaget's (1958) Pendulum problem), or **physical science** (balance beam) areas.

The second sets of instruments were from the social subdomain. Sequences of vignettes (series of similar stories following a common theme) from the Counseling, Child Caregiving Empathy, Helping, and Romantic Breakup sequences were used.

Sequences of vignettes used for Social Subdomain were Counseling, Child Caregiving Empathy, Helping, and Romantic Breakup. Each vignette sequence was followed by questions. For Laundry, Algebra, Infinity and Balance Beam instrument, sequence of mathematical and logic based problems were used. There were a total of 430 item represented in all the instruments together.

Each vignette sequence and problem sequence was at a different order of hierarchical complexity. These included:

1. Preoperational
2. Primary
3. Concrete
4. Abstract
5. Formal
6. Systematic
7. Metasystematic

» METHOD

Participants

There were a total of 409 participants among all instruments from two subdomains. The number participant for each instrument is shown in Table 2. The Age of the participants ranged from 12 years to 87 years. The mean age and standard deviation was $M = 28.42$ ($SD = 8.56$). Education level of the participants varied from elementary school graduate to graduate school; $M = 4$ ($SD = 1.27$). The mean of 4 represents a high school degree

Instruments

Two sets of instruments were used to assess participants' developmental stage. The first sets of instruments were from the logic/mathematics/

Table 2. number of participants for each instrument of the social subdomain and mathematical/ logical/ physical science subdomain

instrument	number of participants
helper person	40
counselor patient	30
empathy	22
breakup dilemma	21
caregiver	42
laundry	111
algebra	40
balance beam	51
infinity	52

Instrument description for social subdomain

1. Empathy: a person explains to another person how to understand the feelings of others in all the vignettes in this sequence
2. Counselor patient: a counselor explains treatment options to a patient in all the vignettes in this sequence
3. Breakup: a person reacts to breaking up with their partner in all the vignettes in this sequence
4. Helper person: a person gives another person guidance and assistance in all the vignettes in this sequence
5. Caregiver: in this sequence of vignettes a caregiver comforts a crying child in all the vignettes in this sequence

An example of social subdomain vignette sequence and corresponding questions is shown in Appendix J.

Instrument description for logic/mathematics/ physical sciences subdomain

1. Algebra (mathematics): this instrument sequence asked participants solve standard algebraic problems

2. Balance beam (physics): this instrument sequence was derived from inhelder and Piaget (1958) balance beam task

3. Infinity (mathematics): this instrument sequence asked participants to solve problems that dealt with conceptualizing infinity

4. Laundry (logic and mathematics): this instrument sequence was a version of Inhelder & Piaget (1958) pendulum task. It asked participants to detect causal relationships from various systems and then compare the systems using logic.

An example of logic/mathematics/physical sciences subdomain problem sequence is shown in Appendix κ.

» PROCEDURE

Participants were asked to rate their preferences for vignette based questions for instruments from the social subdomain. 1–6 rating scale was used as answers to vignette based questions. For instruments from the logic/mathematics/physical sciences subdomain were asked to solve mathematical and logic based problems. The instruments from Social subdomain and logic/mathematics/physical sciences subdomain were given in pairs. Different combination of instruments was used. There were a total of 430 item represented in all the instruments together

Table 3. principal axis factoring with the person Rasch scores for each instrument

| factor | total | initial eigenvalues | |
		% of variance	cumulative %
1	8.146	90.51	90.510
2	0.445	4.947	95.456
3	0.190	2.114	97.570
4	0.129	1.431	99.002
5	0.034	0.376	99.378
6	0.026	0.285	99.663
7	0.016	0.175	99.837
8	0.010	0.108	99.946
9	0.005	0.054	100.00

» RESULTS

There were three parts to the analysis. First, Rasch analysis of all the items from all the instruments was run. This is a psychometric analysis of the responses to the items showing their relative difficulty. The Rasch analysis confirmed that the items were performing as predicted by the MHC. Rasch figures are included in the Appendix so that one can see the actual raw Rasch scores for each item. The Rasch scores showed the raw magnitude of difficulty for each item. Through close inspection of the Rasch map, one can see which items are out of order.

SCREE PLOT

Figure 2. This figure shows the scree plot. The scree plot also showed that just one factor mattered

The Rasch maps in the appendix show that metasystematic items generally had the highest Rasch item scores (RIS) and preoperational items generally had the lowest Rasch item scores. This shows that metasystematic items were the most difficult and preoperational items the most easiest as predicted by MHC. There was some intermixing of items from adjacent stage. For example, in the counselor patient Rasch map, Primary stage items had a higher raw Rasch score in comparison to preoperational stage items. Similarly, formal stage items had a higher raw Rasch score in comparison to systematic stage items. In Infinity Rasch map, some systematic stage items had a higher raw Rasch scores in comparison to the metasystematic stage item. However, intermixing is not too terrible as shown by the high *r*'s for each instrument. The *r*'s for the instrument ranged from .711 to .980. Compared to other instruments, these are very high *r* values.

Second, we report regression analysis assessing how well the order of hierarchical complexity (OHC) of the items predicted the difficulty of the items in terms of their Rasch scores. Because the items have orders of hierarchical complexity (OHC), the relationship between the OHC and the Rasch item scores is psychophysical. There is an analytic property of a stimulus predicting a scaled response. A linear regression shows that order of hierarchical complexity (OHC) of the items predicts the actual difficulty of the items as represented by Rasch items scores. The linear regression showed that hierarchical complexity strongly predicted the stage of performance of all the items from all of the instruments; $r(430) = 0.959$, $p < 0.0001$.

Third, a principal axis factoring was carried out with the person Rasch scores for each instrument. Irrespective of domains, if each instrument loaded on the first factor with all the factor scores over .7 and if the first factor accounted for more than 70% of the variance, then that would show that all instruments were part of a single domain. The analysis showed that irrespective of domains, the person scores on each instrument loaded on the first factor (90.510% of the variance). There were only very low loadings on the second factor (4.947% of the variance) as shown in Table 2. The scree plot also showed that only the first factor mattered. This is shown in Figure 2. All the factor scores on the first factor were over .85 as shown in Table 4.

In each case the MHC accounted for the largest amount of variance with *r* values of $r(40) = .977$, $p \leq .001$ for helper person, $r(30) = .934$, $p \leq .001)$ for counselor patient, $r(22) = .910$, $p \leq .001)$ for empathy, $r(21) = .835$, $p \leq .001$ for breakup dilemma, $r(42) = .711$, $p \leq .001$ for caregiver, $r(51) = .980$, $p \leq .001$ for balance beam, $r(40) = .966$, $p \leq .001$ for algebra, $r(111) = .964$, $p \leq .001$ for laundry, and $r(52) = .912$, $p \leq .001$ for the infinity instrument. Rasch map and result of stepwise regression for all the instruments is shown in the appendix section.

The critical test for determining if there is just a single domain is to conduct a factor analysis. The first factor accounted for 90.510% of the variance. The second factor accounted for only

4.947% of the variance. This showed only one factor mattered. That shows all the instruments are part of a single domain

Table 4 shows the factor scores of all the instruments in the first factor and second factor. For missing values, cases were excluded listwise. All the factor scores on the first factor were over .85. Overall factor scores show the presence of just one factor. The second factor does not show in the factor matrix.

» DISCUSSION

MHC assumes axiomatically that it is content and context free. A predicted consequence of this is that stage of performance on tasks of a given order of hierarchical complexity (OHC) should not depend on the content and context of the items in an assessment. This prediction was born out in the present study.

Here, order of hierarchical complexity (OHC) of the items almost perfectly predicted Rasch performance scores that reflected actual performance-measured difficulty among these instruments. The outcome was that the factor analysis found only a single factor. This implies that the instruments in social subdomain and instruments in the logic/mathematics/physical sciences subdomain belong to a single domain. This may explain why Inhelder & Piaget (1958) thought that development was synchronous across all tasks and domains.

Consider the two tasks presented next. They differ in content, form and method. The content and form, method of assessment of the helper-person task is entirely different from the content, form and method of assessment in the laundry task. The helper-person task has social content, a series of vignettes that are assessed by having the participants rate how well the method of interacting

Table 4. factor matrix

instruments	factor	
	1	2
helper person	0.973	0.129
infinity	0.968	0.142
counselor patient	0.964	–0.060
caregiver	0.962	–0.174
depression breakup	0.962	–0.181
algebra	0.952	0.093
empathy	0.942	–0.102
laundry	0.934	–0.271
balance beam	0.873	0.455

note: extraction method: principal axis factoring

with the person to be helped (method). The laundry problem has chemical content. It presents how various ingratiated and washing methods might make a stain in cloth come out clean or dirty. They make predictions that are right or wrong. Yet they both load extremely highly on a single factor and the order of hierarchical complexity predicts the difficulty of performance extremely well.

What is the status of content and context claims about differences in stage given their lack of empirical support found in our study. Maybe what other views of domains reflect, is that content analysis is possible. One can analyze the content of a task and the method of presentation, but that does not mean that the content or method make a difference in stage. But we found content and form of the tasks to be unrelated to the stage of performance. That does not mean there are not people with performance deficits in given domains. These performance deficits may be because of lack of practice due to a lack of interest in the domain or lack of exposure to tasks in the domain or in some cases specific deficits results in unevenness in performance.

Content analysis has some use in describing the differences in the type of tasks to be completed. This does not mean that the task demands are identical or tasks actions are all the same. It just says that OHC required by different tasks is what is telling. In the small effects paper (Commons, Giri & Harrigan, 2014), we give further evidence that although there are many sources of differences, they are all overwhelmed by the predictive power of OHC.

We conclude there is only one stage domain. In the future, we plan to see what the relationship is between this domain and other measures of "smarts" such as IQ. ■

REFERENCES

Andrich, D., & Constable, E. (1984, March). Studying unfolding developmental stage data based on Piagetian tasks with a formal probabilistic model. Paper presented at the Annual meeting of the AERA, New Orleans.

Bond, T. G. (1994). Piaget and measurement II: Empirical validation of the Piagetian model. Archives de Psychologie, 63, 155–185.

Bond, T. G., & Fox, C. M. (2001). Applying the Rasch model: Fundamental measurement for the human sciences. Mahwah, NJ: Lawrence Erlbaum Associates, Inc

Case, R. (1991). The mind's staircase: Exploring the conceptual underpinnings of children's thought and knowledge. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Case, R., Okamoto, Y., Henderson, B., & McKeough, A. (1993). Individual variability and consistency in cognitive development: New evidence for the existence of central conceptual structures. In R. Case & W. Edelstein (Eds.), The new structuralism in cognitive development: Theory and research on individual pathways (pp. 71 − 100). Basel, Switzerland: Karger.

Commons, M. L., Armon, C., Richards, F. A., Schrader, D. E., Farrell, E. W., Tappan, M. B., & Bauer, N. F. (1989). A multidomain study of adult development. In D. Sinnott, F. A. Richards, & C. Armon (Eds.), Adult development, Vol. 1: Comparisons and applications of developmental models. (pp. 33 − 56). New York, NY: Praeger

Commons, M. L., Trudeau, E. J., Stein, S. A., Richards, S. A., & Krause, S. R. (1998). Hierarchical complexity of tasks shows the existence of developmental stages. Developmental Review, 18, 237 − 278.

Commons, M.L., & Richards, F. A. (2003). Four Postformal Stages. In Demick, J. & Andreoletti, C. (Eds.). Handbook of adult development (pp. 199–219), New York: Plenum Press.

Commons, M. L., & Pekker, A. (2008). Presenting the formal theory of hierarchical complexity. World Futures: Journal of General Evolution, 64(5−7), 375–382. doi:10.1080/02604020802301204

Commons, M. L., Giri, S. P., & Harrigan, W. H. (2014). The small effects of non-hierarchical complexity variables on performance. In Press.

Dawson, T. L. (1998). ''A good education is.'' A lifespan investigation of developmental and conceptual features of evaluative reasoning about education. Unpublished doctoral dissertation, University of California at Berkeley, Berkeley, CA.

Dawson, T. L. (2000). Moral reasoning and evaluative reasoning about the good life. Journal of Applied Measurement, 1(4), 372 − 397.

Dawson, T. L. (2002A). A comparison of three developmental stage scoring systems. Journal of Applied Measurement, 3, 146 − 189.

Dawson, T. L. (2002b, January). Measuring intellectual development across the lifespan. Paper presented at the powerful learning & the Perry scheme: Exploring intellectual development's role in knowing, learning, and reasoning, California State University, Fullerton, CA.

Dawson, T. L. (2002c). New tools, new insights: Kohlberg's moral reasoning stages revisited. International Journal of Behavioral Development, 26, 154 − 166.

Dawson, T. L., & Gabrielian, S. (2003). Developing conceptions of authority and contract across the lifespan: Two perspectives. Developmental Review, 23, 162 − 218.

Dawson, T. L., Xie, Y., & Wilson, M. (2003). Domain-general and domain-specific developmental assessments: Do they measure the same thing? Cognitive Development, 18, 61 − 78.

Dawson-Tunik, T. L. (2004a). ''A good education is . . .'' The development of evaluative thought across the lifespan. Genetic, Social, and General Psychology Monographs, 130(1), 4 − 112.

Dawson, T. L. (2004). Assessing intellectual development: Three approaches, one sequence. Journal of Adult Development, 11, 71 − 85.

Demetriou, A., & Valanides, N. (1998). A three-level theory of the developing mind: Basic principles and implications for instruction and assessment. In R. J. Sternberg & W. M. Williams (Eds.), Intelligence, instruction, and assessment: Theory into practice (pp. 211 − 246). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Draney, K. L. (1996). The polytomous Saltus model: A mixture model approach to the diagnosis of developmental differences. Unpublished doctoral dissertation, University of California at Berkeley, Berkeley, CA.

Erdynast, A., Armon, C., & Nelson, J. (1978). Cognitive developmental conceptions of the true, the good, and the beautiful. The eighth annual proceedings of Piaget and the helping professions. Los Angeles: University of Southern California Press.

Erdynast, A., & Chen, W. (2014). Relations between Adult Developmental Conceptions of the Beautiful and Moral Development. In Press.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. Psychological Review, 87, 477–531.

Fischer, K. W., & Bidell, T. R. (1998). Dynamic development of psychological structures in action and thought. In W. Damon & R. M. Lerner (Eds.), Handbook of child psychology: Theoretical models of human development (5th ed., pp. 467 − 561). New York, NY: Wiley.

Inhelder, B., & Piaget, J. (1958). The Growth of Logical Thinking from Childhood to Adolescence. Basic Books, New York, NY.

Kegan, R. G. (1979). The Evolving Self: A Process Conception for Ego Psychology. The Counseling Psychologist, 8, 2, 5–34.

King, P. M., Kitchener, K. S., Wood, P. K., & Davison, M. L. (1989). Relationships across developmental domains: A longitudinal study of intellectual, moral, and ego development. In M. L. Commons, J. D. Sinnot, F. A. Richards, & C. Armon (Eds.), Adult development. Volume 1: Comparisons and applications of developmental models (pp. 57 − 71). New York, NY: Praeger.

Kitchener, K. S., Lynch, C. L., Fischer, K. W., & Wood, P. K. (1993). Developmental range of reflective judgment: The effect of contextual support and practice on developmental stage. Developmental Psychology, 29, 893 − 906.

Kohlberg, L., Levine, C., & Hewer, A. (1983). Moral stages: A current formulation and a response to critics. (pp. 63–128).Basel: Karger.

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. Goslin (Ed.), Handbook of socialization theory and research (pp. 347 − 480). Chicago, IL: Rand McNally.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). Foundations of measurement, Vol. I: Additive and polynomial representations. (pp. 93–112). New York: Academic Press.

Larivee, S., Normandeau, S., & Parent, S. (2000). The French connection: Contributions of French-language research in the post-Piagetian era. Child Development, 71(4), 823 − 839.

Luce, R. D. & Tukey, J. W. (1964). Simultaneous conjoint measurement: a new scale type of fundamental measurement. Journal of Mathematical Psychology, 1(1), 1−27. doi:10.1016/0022−2496(64)90015-X

Müller, U., Sokol, B., & Overton, W. F. (1999). Developmental sequences in class reasoning and propositional reasoning. Journal of Experimental Child Psychology, 74, 69 − 106.

Pascual-Leone, J., & Goodman, D. (1979). Intelligence and experience: A neoPiagetian approach. Instructional Science, 8, 301 − 367.

Perry, W.G. (1968). Forms of intellectual and ethical development in the college years: A scheme. New York: Holt, Rinehart & Winston.

Piaget, J. (2000). Studies in reflecting abstraction (R. L. Campbell, Trans.). Hove, UK: Psychology Press.

Rasch, G. (1980). Probabilistic model for some intelligence and attainment tests. Chicago, IL: University of Chicago Press.

Shannon, C. E., & Weaver, Warren. (1963). The Mathematical Theory of Communication. Urbana, IL: University of Illinois Press.

Smith, R. M. (Ed.) (2004). Introduction to Rasch measurement. Maple Grove, MN: JAM Press

Turiel, E. (1980). The development of social-conventional and moral concepts. In M.

Windmiller, N. Lambert, & E. Turiel (Eds.), Moral development and socialization (pp. 69 − 106). Boston, MA: Allyn & Bacon.

Wilson, M. (1984). A psychometric model of hierarchical development. Unpublished doctoral dissertation, University of Chicago, Chicago, IL, USA.

Wilson, M. (2005). Constructing measures: An item response modeling approach. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

APPENDIX

» APPENDIX A1
Helper person rasch map

```
                                    pre  = Preoperational
                                    pri  = Primary
                                    conc = Concrete
        Higher Stage                abs  = Abstract
2            +                       fl   = Formal
             |                       sys  = Systematic
             |T                      me   = Metasystematic
             |
        X    |  me1.1    me1.2
        X    |  me1.3
          T  |  me1.4
             |  me1.5    sys1.2
1            +S
        XX   |  f1.2     sys1.1
        X    |  sys1.3   sys1.4   sys1.6
          S  |  f1.1     f1.4     sys1.5   sys1.7
        X    |  f1.5
 XXXXXXXX    |  abs1.1   abs1.2   abs1.4   f1.3      f1.7
   XXXXXX    |  abs1.3   f1.6
      XXX  M|M abs1.6    abs1.7
0            +  abs1.5
     XXXX    |
       XX    |
        X    |
   XXXXXX  S|  pri1.1
        X    |S pri1.2
        X    |  pri1.3   pri1.4
             |  pre1.1   pre1.3   pri1.5   pri1.6   pri1.7
-1      X  T+  pre1.2   pre1.4   pre1.5   pre1.6   pre1.7
        X    |
             |
             |
             |T
             |
             |
-2           +
        Lower Stage
```

There were 40 participants ranging in age from 19 to 63 ($M = 22.16$, $SD = 9.09$) with education varying from elementary school to graduate school. Each × in the Rasch map represents a participant. In this figure, the concrete stage items are missing because; the items did not perform as designed.

» APPENDIX B1
Counselor patient rasch map

```
                                    pre  = Preoperational
                                    pri  = Primary
                                    conc = Concrete
        Higher Stage                abs  = Abstract
2           +T                       fl   = Formal
            |                        sys  = Systematic
       X    |  me1.3                 me   = Metasystematic
            |  me1.2
         T  |  me1.1    me1.4    me1.5    me1.6
            |  me1.7
            |  fl1.1    fl1.2    fl1.6
1         +S   fl1.3    fl1.4
     XX     |  fl1.5
   XXX  S  |  fl1.7
     X     |  sys1.1   sys1.2
   XXX     |  sys1.4   sys1.5   sys1.6
           |  sys1.3   sys1.7
           |
0    XX    +
   XXX  M|M abs1.6
 XXXXX     |  abs1.1   abs1.2   abs1.3   abs1.5   abs1.7
    XX     |  abs1.4
 XXXXX     |  cr1.3    cr1.4    cr1.6
           |  cr1.1    cr1.2    cr1.5
        S  |  cr1.7
-1         +
           |  pre1.4
           |S pre1.1   pre1.2   pre1.6
     X     |  pre1.3   pre1.5   pre1.7
   XX   T  |  pri1.2   pri1.4   pri1.5
           |  pri1.1   pri1.3   pri1.6   pri1.7
           |
-2         +
        Lower Stage
```

There were 30 participants ranging in age from 16 to 71 ($M = 25.13$, $SD = 7.02$) with education varying from high school to graduate school. Each × in the Rasch map represents a participant

» APPENDIX A2
Helper person results

Scatter plot of Rasch score versus order of hierarchical complexity. Legend — OHC item: 7 preoperational, 8 primary, 9 concrete, 10 abstract, 11 formal, 12 systematic, 13 metasystematic. R^2 linear = 0.955, $r(40) = 0.977$
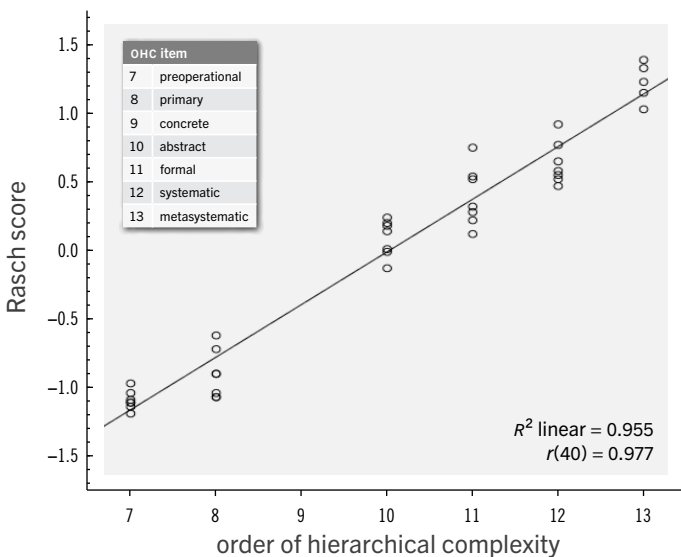
» APPENDIX B2
Counselor patient results

Scatter plot of Rasch score versus order of hierarchical complexity. Legend — OHC item: 7 preoperational, 8 primary, 9 concrete, 10 abstract, 11 formal, 12 systematic, 13 metasystematic. R^2 linear = 0.873, $r(30) = 0.934$
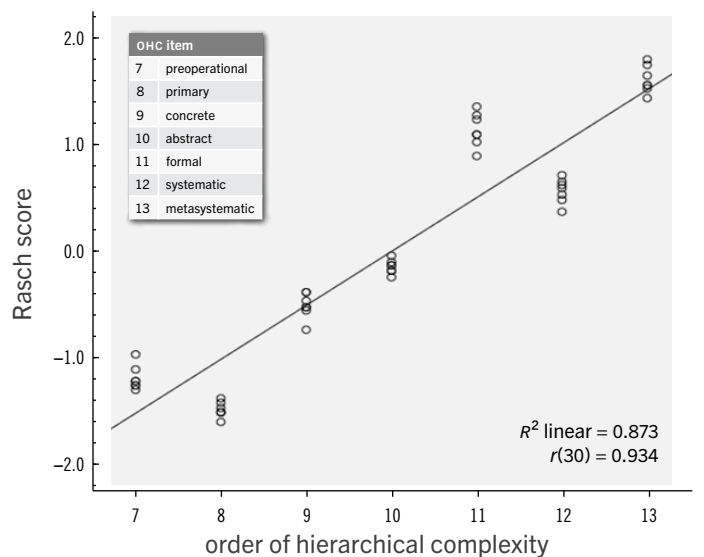
» APPENDIX C1
Empathy rasch map

```
                        Higher Stage
  2        +                                          :
        X  |                                           :
           |                                           :
           |T                                          :
        T| 
           |    me1.4    me1.5    me1.6                 :
           |    me1.1    me1.2    me1.3    me1.7    me  *
           |  f1.5
           |  f1.2    f1.4
        X  |S f1.6    f1.7
  1        S+ sys1.2   sys1.3   sys1.6
        X  |  f1.1    f1.3      sys1.1   sys1.4
           |  sys1.5   sys1.7
     XXX   |
      XX   |
     XXX   |M
        X  |  abs1.1   abs1.2   abs1.3
      X M|  abs1.4   abs1.5   abs1.7   pre1.1
        X  |  cr1.1    cr1.4    cr1.6   pre1.3
           |  abs1.6   cr1.2    cr1.3   cr1.5    cr1.7
  0     X  +  pre1.2   pre1.6
        X  |S pre1.5   pre1.7
        X  |  pre1.4
           |  pri1.1   pri1.2   pri1.3   pri1.6
      X S|  pri1.4   pri1.5   pri1.7
        X  |
           |
           |T
       XX  |
 -1        +
        T|
        X  |
           |                            pre  = Preoperational
           |                            pri  = Primary
           |                            conc = Concrete
           |                            abs  = Abstract
           |                            fl   = Formal
           |                            sys  = Systematic
 -2        +                            me   = Metasystematic
       Lower Stage
```

There were 22 participants ranging in age from 18 to 65 (M = 9.00, SD = 2.02) with education varying from high school to graduate school. Each × in the Rasch map represents a participant
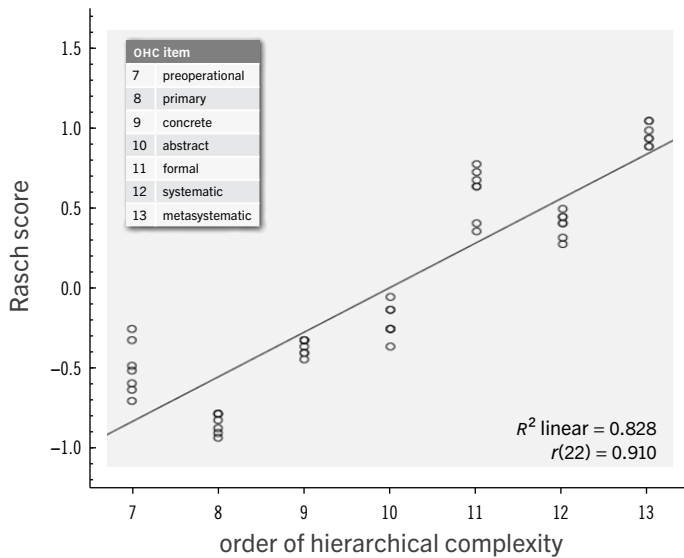
» APPENDIX D1
Depression breakup dilemma rasch map

```
                      Higher Stage
  1             +            pre  = Preoperational
                |            pri  = Primary
                |            conc = Concrete
                |            abs  = Abstract
          X     |            fl   = Formal
                |            sys  = Systematic
  0      XX  S+             me   = Metasystematic
          X     |
          X     |
          X     |
          X     |
         XX     |  me1.2
         XX  |T me1.1
          X  |  me1.4
 -1      XXX  +
          M|  sys1.2
           |  sys1.1   sys1.4
          X  |  me1.3
           |
           |S
           |  sys1.3   sys1.5
         XX  |  me1.5
 -2         +
           |M cr1.5    pri1.5
          S|  abs1.5   f11.5    pri1.2
           |  f11.1    f11.2    f11.3   pri1.1
           |  abs1.2   f11.4    pri1.3
           |  abs1.4   pri1.4
          X  |  abs1.1   abs1.3   cr1.3   pre1.2
           |  pre1.3
          |S cr1.2
 -3         +  cr1.1    cr1.4
           |  pre1.4   pre1.5
          X  |
          T|  pre1.1
           |
           |T
           |
 -4         +
           |
           |
           |
          X  |
            +
       Lower Stage
```

There were 21 participants ranging in age from 18 to 65 (M = 24.07, SD = 9.02) with education varying from elementary school to graduate school. Each × in the Rasch map represents a participant

» APPENDIX C2
Empathy results

OHC item

7	preoperational
8	primary
9	concrete
10	abstract
11	formal
12	systematic
13	metasystematic

R^2 linear = 0.828
$r(22) = 0.910$

» APPENDIX D2
Depression breakup dilemma results

OHC item

7	preoperational
8	primary
9	concrete
10	abstract
11	formal
12	systematic
13	metasystematic

R^2 linear = 0.697
$r(21) = 0.835$
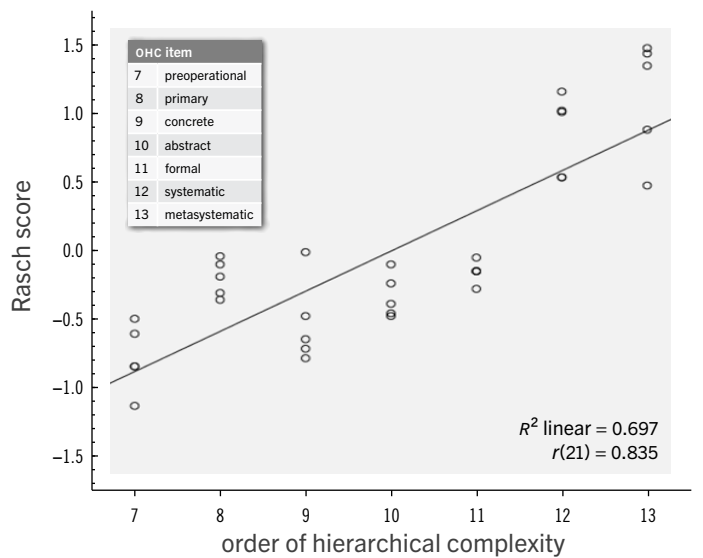
» APPENDIX E1
Caregiver rasch map

```
Higher Stage
     4        +
              |
              |
       XX     |
              |
     3        +
              |
              |
         T|
              |
     2        +
              |
       X  |
    XXXX S|
       X  |
       X  |
     1  XXX  +T
       X  |    Me.2  Me.3
       XX  |
      XXX  |    Me.1  Me.4  Me.5
      XXX M|S Conc.2  Conc.3  Fl.2  Fl.3
     XXXX  |  Conc.1  Fl.1  Fl.5  Sys.2  Sys.3
       XX  |  Conc.4  Conc.5  Fl.4  Sys.1
     0 XXXX  +M Abs.1  Abs.3  Sys.4  Sys.5
       XX  |
       X  |    Abs.4  Pre.1  Pre.2  Pre.3
       X  |S Abs.5  Pre.4  Pre.5
      XX S|  Abs.2  Pri.3
      XX  |  Pri.1
           |  Pri.2
    -1   X  +T Pri.4  Pri.5
              |
              |
              |
         T|
       X  |
       X  |
    -2        +
          Lower Stage
```
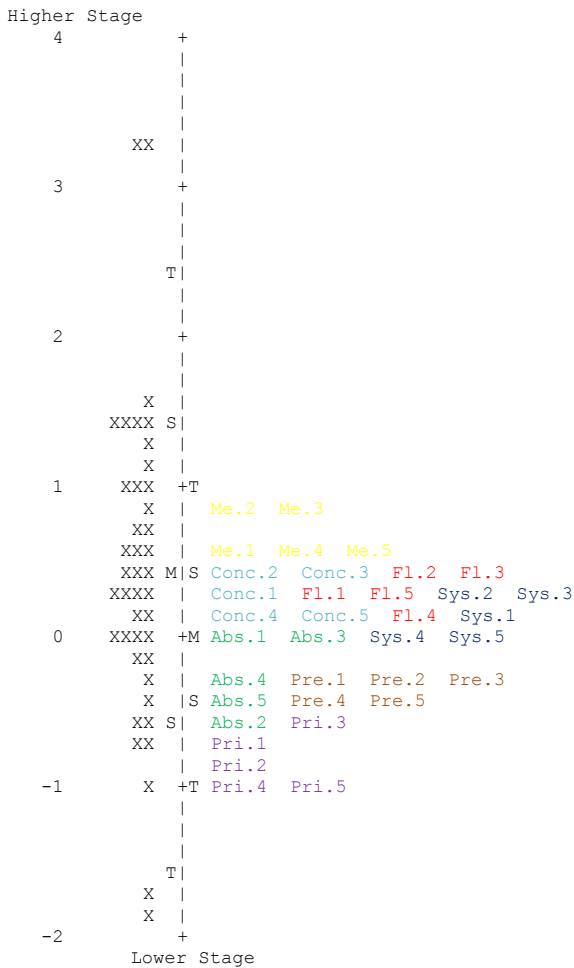
There were 42 participants ranging in age from 19 to 62 (M = 22.90, SD = 8.12) with education varying from elementary school to graduate school. Each × in the Rasch map represents a participant

» APPENDIX E2
Caregiver results
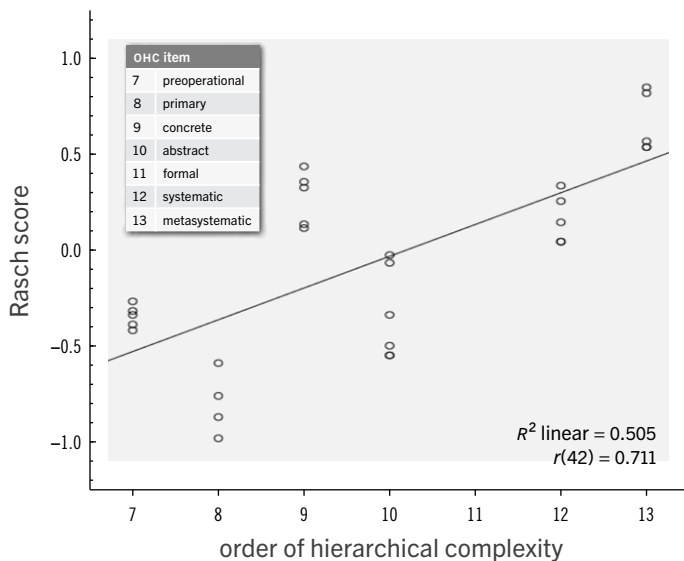
» APPENDIX F1
Laundry rasch map

```
          Higher Stage
     8        +  me  me
              |
         T|  me  me  me
     7        +
          |T  me  me
              |  me
     6        +  me
              |  me
       XXX  |  sys
     5 XXXXX  +
       XXX  |
     XXXXXX  |  sys
       X  |  sys
     4 XXXXXXX S+  sys
      XXXX  |  sys  sys
          |S sys
       XXX  |
       X  |  sys  sys  sys
     3   X  |  sys  sys  sys  sys  sys
       XXX  |  sys  sys  sys  sys  sys  sys
     XXXXXX  |  sys  sys  sys  sys  sys
       XXX  |  Sys  sys
     2 XXXX  |  sys  sys
      XXXX  |  fl
      XXXX  |  fl  fl  fl  fl  fl  fl  fl  fl  fl  fl  fl  fl
       X M+  fl  fl  fl  fl  fl  fl
     1        |  fl
       X  |
       XX  |
     0        +M
    XXXXXXXXXXX  |
       XXX  |
       X  +
    -1       X  |  abs
       X  |  abs  abs  abs  abs  abs  abs  abs  abs  abs  abs
       +  abs  abs  abs  abs
    -2       S|  abs  abs  abs  abs  cr
      XXX  |  cr  cr  cr
       X  |  cr  cr  cr
       +  cr  cr  cr  cr
    -3       X  |  cr
          |S
    XXXXXX  |  pri
    -4 XXXXXXX  +  pri  pri  pri
       XXX  |
       X  |  pri  pri  pri  pri  pri  pri
       X  |  pri  pri  pri  pri
    -5   X  +  pri  pri
          |  pri  pri  pri
         T|  pri
    -6        +
          Lower Stage
```

There were 111 participants ranging in age from 12 to 73 (M = 26.02, SD = 10.62) with education varying from elementary school to graduate school. Each × in the Rasch map represents a participant

» APPENDIX F2
Laundry results
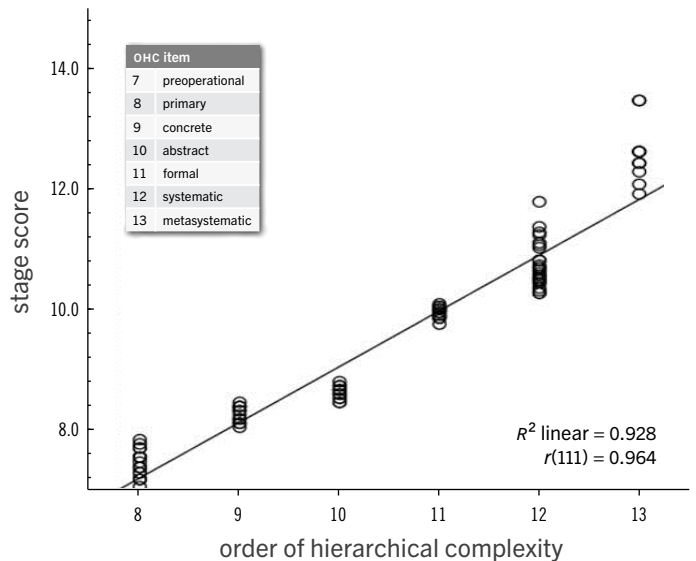
» APPENDIX G1
Algebra rasch map

» APPENDIX H1
Balance beam rasch map
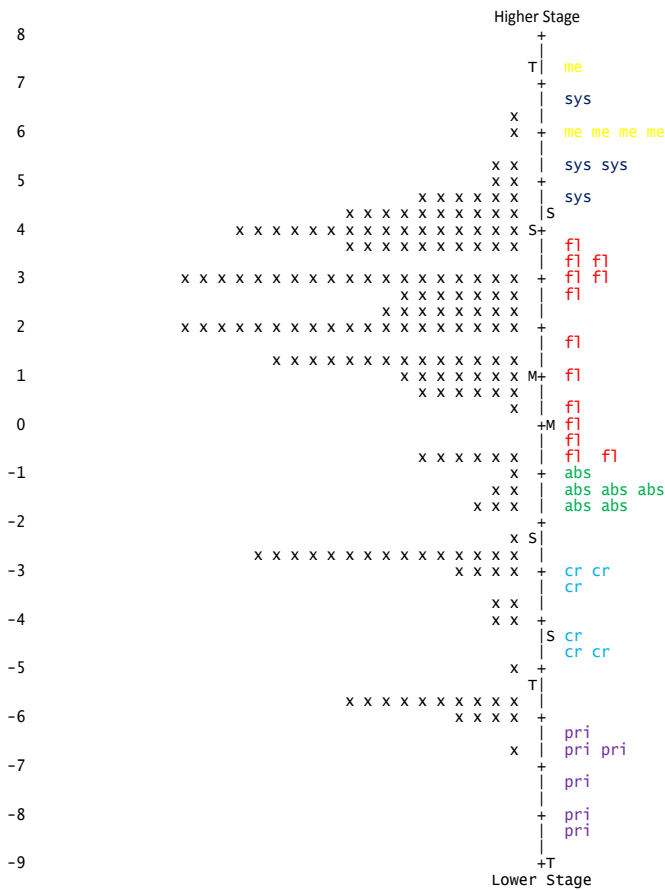
There were 40 participants ranging in age from 14 to 81 (M = 24.06, SD = 8.60) with education varying from elementary school to graduate school. Each × in the Rasch map represents a participant.
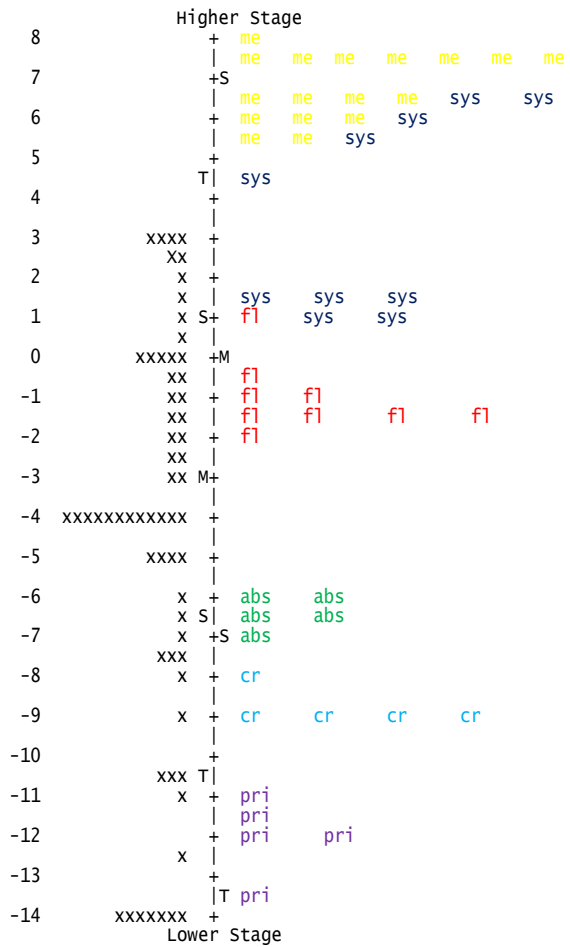
There were 51 participants ranging in age from 12 to 87 (M = 24.11, SD = 9.05) with education varying from elementary school to graduate school. Each × in the Rasch map represents a participant.

» APPENDIX G2
Algebra results
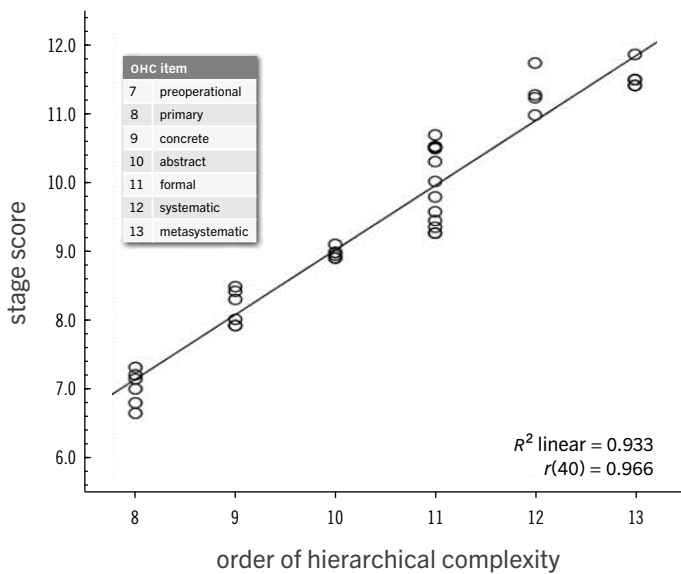
» APPENDIX H2
Balance beam results
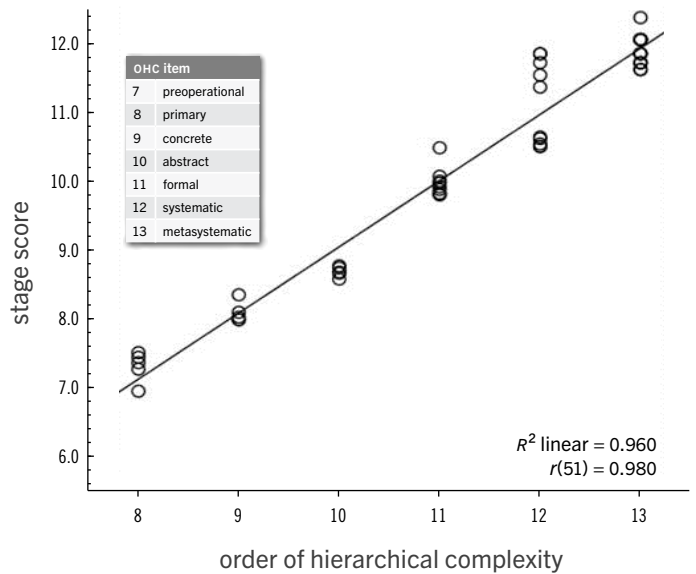
» APPENDIX I1
Infinity rasch map
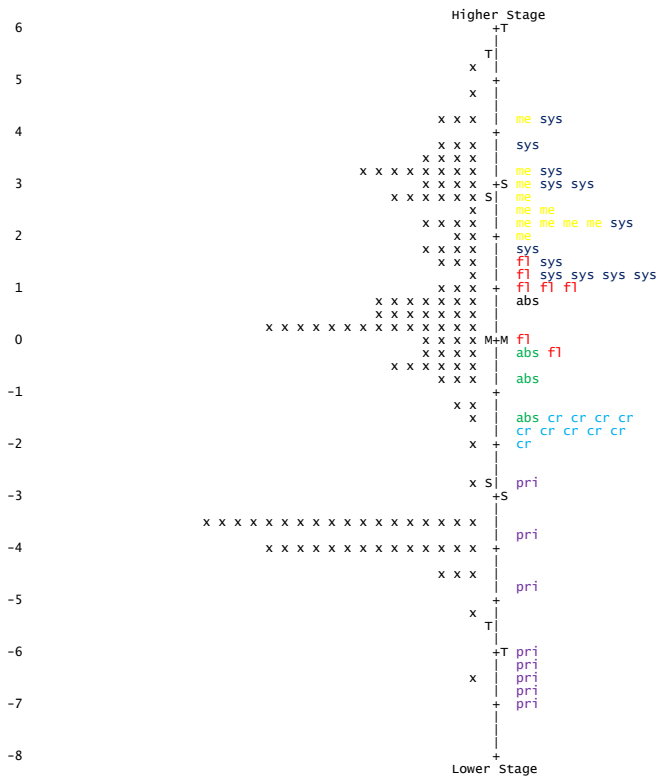
There were 51 participants ranging in age from 12 to 73 (M = 25.05, SD = 10.51) with education varying from elementary school to graduate school. Each × in the Rasch map represents a participant.

» APPENDIX I2
Infinity results

OHC	item
7	preoperational
8	primary
9	concrete
10	abstract
11	formal
12	systematic
13	metasystematic

R^2 linear = 0.831
$r(52)$ = 0.912
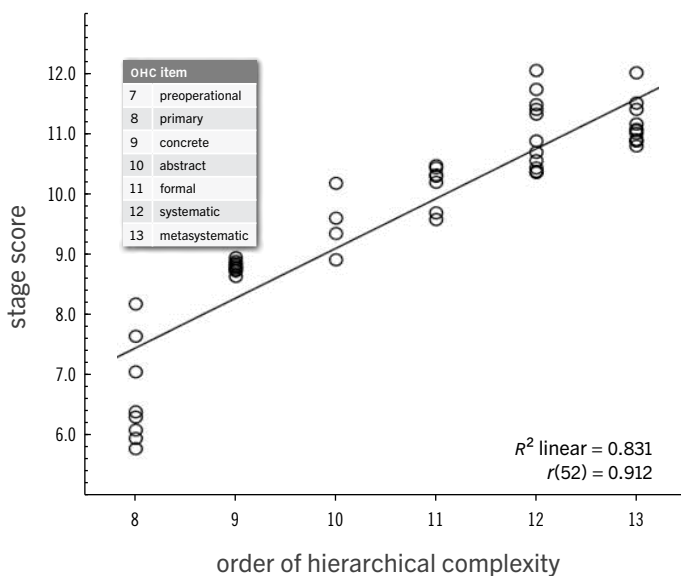
» APPENDIX J
Social subdomain vignette sequence and questions

Helper person metasystematic vignette example with questions
White speaks with the Person to assess the problem. During the conversation, White offers to provide guidance and assistance seen as most effective in treating this problem. White presents other forms of guidance and assistance as well, and discusses the benefits and risks of each, including doing nothing. White, seeking to understand the Person's needs and concerns, asking and answering many questions. White also sees if the Person's body language matches their statements. White asks if the Person is ready to make a choice based on their previous discussion. Feeling White knows best, the Person accepts the guidance and assistance.

1. Based on the paragraph above, rate each question by selecting a number on the following scales. A rating of 1 means you strongly disagree with Helper. A rating of 6 means you strongly agree with the Helper.

	Extremely poor				Extremely good	
	1	2	3	4	5	6
Rate White's method of offering guidance and assistance.	○	○	○	○	○	○
Rate how clearly White expressed their idea.	○	○	○	○	○	○
Rate the degree to which White informed their Person.	○	○	○	○	○	○

2. Rate these two questions on a 1–6 point scale. 1 means you are not at all likely to hire or recommend this Helper, 6 means you are extremely likely to hire or recommend this Helper.

	Not at all likely				Extremely likely	
	1	2	3	4	5	6
Rate how likely you would be to accept the guidance and assistance offered by White.	○	○	○	○	○	○
Rate how likely you would recommend White's guidance and assistance.	○	○	○	○	○	○

» APPENDIX J
Logic/mathematics/physical sciences subdomain problem sequence
Balance beam metasystematic problem sequence example

Subsystem A: In one class, $x + 6$ students each received a rating of 1. The other 5 students received a rating of x.

ratings	1	x
number of students	$x + 6$	5

Subsystem B: In one class, $y + 6$ students each received a rating of 500. The other 5 students received a rating of $500y$.

ratings	500	$500y$
number of students	$y + 6$	5

**BEHAVIORAL
DEVELOPMENT
BULLETIN**

Relations between adult developmental conceptions of the beautiful and moral development

Albert Erdynast and Wendy Chen

Antioch University

The cross-sectional study with 180 adult subjects researched structural-developmental levels of conceptions of the beautiful and also studied their relations with structural-developmental levels of moral development. Measures of moral development that are considerably beyond the scope of Kohlberg's conception of justice-reasoning were used. Six levels of conceptions of the beautiful were studied. The highest level, not discerned in previous studies, was established from some responses to Picasso's Les Demoiselles d' Avignon. Six structural-developmental conceptions of the beautiful across four Picasso paintings and one Michelangelo sculpture meet the criterion of generality which supports the claim to finding hierarchical structures within a domain of their own. Distinctions are made about the domain of the issues under study among four domains: I) the real, II) the good, III) the just and right, and, IV) the beautiful. The beautiful, domain IV, involves the contemplation, or, appreciation, and fashioning of beautiful objects. Data were collected and analyzed in two domains and two sub-domains of moral development characterized as: 1) conceptions of the good, 2) judgments of justice-reasoning (obligations, liberties and duties), and 3) compassion as a supererogatory act which a person does for the sake of another's good at considerable cost or risk to the self. The moral domains and subdomains used as a framework for the measures of moral development in these studies are based on the framework of Rawls's social contract philosophy and the psychological study of conceptions of the beautiful is based on Kant's philosophy of the beautiful. The relations between the developmental level of conceptions of the beautiful and the structural-developmental levels of moral decisions were studied through standardized justice-reasoning dilemmas and administration of a fidelity/infidelity structural interview. The data suggest that as the level of moral development rises, there is a tendency for the levels of conceptions of the beautiful to also rise. In 91% of instances, there was plus or minus a half level correspondence between the level of conceptions of justice-reasoning and the level of conceptions of the beautiful. In 98% of the analyzed cases, the levels of conceptions of the good and the levels of conceptions of the beautiful are within one level of one another. The study also contributes toward resolving the unsettled question about the number of developmental domains and subdomains that exist.

KEYWORDS: moral development, beautiful, adult development

STRUCTURAL-DEVELOPMENTAL CONCEPTIONS OF the beautiful have not been previously established as general hierarchical structures within a domain of their own. Also, the number of developmental domains and subdomains that exist is an unsettled question. This study addresses both sets of issues. The existence of a distinct domain of the beautiful and the issue of the number of domains and subdomains into which to organize developmental theory has considerable import on central concepts that are used, types of research methodologies that are effective in structuring studies, organizing findings and deriving pedagogic practices

that stimulate further development. If the number of domains is plural, they cannot be conflated into just a single one. In the domain of factual reality, claims are proven or not proven, or, facts are inferred from evidence. Moral judgments and judgments of the beautiful do not derive from proofs. Principled conceptions of the good are relative to the individual's aims, ends, worthwhile interests and attachments to individuals and associations. In the subdomain of justice-reasoning, principles that meet the criteria of generality, universality, ordering, publicity and finality (Rawls, 1971) to resolve competing claims are chosen. In the domain of the beautiful, principled appreciation of works of art govern judgments of the beautiful as these are independent from personal judgments of tastes, likes, and dislikes or cultural relativity.

Author note: correspondence regarding this article should be addressed to Dr. Albert Erdynast and Wendy Chen Antioch University, 400 Corporate Pointe Culver City, CA 90230 E-mail: aerdynast@antioch.edu

» PRINCIPAL OBJECTIVES OF THE STUDY

The principal objectives of the study are:

1. To directly study conceptions of the beautiful and to examine whether there is an existent developmental level that is higher than the five levels (or stages) identified in the Erdynast, Armon and Nelson (1978) study and in the Parsons (1987) study.

2. To study the relations between structural-developmental conceptions of the beautiful and structural-developmental conceptions of moral development.

In the view of Giri, Commons, and Harrigan (2014), there is only one single domain with several subdomains. In the view of Kohlberg and Armon (1984) and Piaget (1965), there are only two domains with "hard structural-developmental stages," which are the stages of factual reality (Piaget, 1965), and those of justice-reasoning (Kohlberg, 1969). These hard stage theories can be viewed in contrast to numerous "soft" stage theories such as those articulated by Erikson (1959), Kegan (1979), Levinson, (1978), Loevinger (1976), Perry (1968) and Selman (1980).

Some theories are single domain theories such as Kohlberg's theory of justice reasoning (Kohlberg, Levine, & Hewer, 1983), and Commons and Richards' theory of factual reality (2003).Other theories are multi-domain theories. Perry's 1968 study *Forms of Intellectual and Ethical Development in the College Years* researched across two domains: epistemological and meta-ethical development. Erdynast, Armon and Nelson's 1978 study *Cognitive-Developmental Conceptions of the True, the Good and the Beautiful* researched across three domains: conceptions of factual reality, conceptions of the good, and conceptions of the beautiful.

Snarey, Kohlberg, and Noam (1982) refer to single domain theories as a mono-domain approach, and refer to multi-domain theories as multi-subdomains. In the view of Erdynast, Armon, and Nelson (1978), and in the conceptual framework of this research study, there are at least four theoretically distinct domains. These four domains are: I) the real, II) the good, III) the just and right and IV) the beautiful (Table 1).

» CONCEPTIONS OF THE BEAUTIFUL

Conceptions of the beautiful, Domain IV, are the contemplation, appreciation, and fashioning of beautiful objects. In the aesthetic development domain of art, the fundamental question (Adler, 1981) is: Is the work of art itself beautiful? The works of Immanuel Kant and Meredith (1911), Baldwin (1915) and Dewey (1934) propose aesthetic principles for analyzing artistic beauty. Empirical studies of developmental levels of conceptions of the beautiful, however, are exceptionally sparse. Studies of Kantian conceptions of the beautiful as an irreducible category of judgment that cannot be conflated with other categories of judgment (Marshall, 1922; Baldwin, 1915) are rare. Piaget and Kohlberg, who conducted extensive cognitive developmental studies, did not include the arts in their studies as noted by Gardner, Winner, and Kircher (1975) who studied children's conceptions of art works. In that cross-sectional study, 121 children ranging in age from four to sixteen were shown works of art and administered an open-ended Piagetian clinical interview with inquiries about reasons for the children's tastes such as, "Did you like it?" and their opinions—"Do you think everyone would like it?" Three general developmental levels of responses were found. These responses were characterized in a quasi-rating system as *immature* responses among the 54, four-to-seven year olds; *intermediate* or *transitional* responses among the 51 eight-to-twelve year olds, and *mature* responses among the 16 fourteen-to-sixteen year olds. Parsons (1987) examined understanding of art and presented over three hundred children, adolescents, and adults with 7 questions about 8 paintings. The questions are about art and aesthetics as they are philosophically conceived (Collingwood, 1958; Danto, 1981; Dewey, 1934; Langer, 1953). The statement and questions included:

1. Describe this painting to me.

2. What is it about? Is that a good subject for a painting?

3. What feelings do you see in the painting?

4. What about the colors? Are they good colors?

5. What about the form (things that repeat)? What about texture?

6. Was this a difficult painting to do? What would be difficult?

7. Is this a good painting? Why?

Parsons' questionnaire, administered to over 300 subjects ranging from pre-school children to art professors, was effective in identifying cross-sectional developmental conceptions of the good as applied to art works. The works of art for the study included:

1. Renoir: *The Luncheon of the Boating Party,* 1881.

2. Ivan Albright: *Into the World Came a Soul Called Ida*, 1930.

3. Picasso: *Guernica*, 1936.

4. Paul Klee: *Head of a Man* (sometimes called *Senecio*), 1922.

5. George Bellows: *Dempsey and Firpo*, 1924.

6. Goya: *Lo Mismo*, from *The Disasters of War*, 1810-1820.

7. Picasso: *Head of Weeping Woman with Hands*, 1936.

8. Chagall: *La Grande Cirque*, 1927.

The stages of understanding art were determined by examining responses about *1)* the subject, *2)* expression, *3)* the medium, form and style, and *4)* judgment. The developmental stages of understanding art identified in the study are: Stage One) Favoritism, Stage Two) Beauty and Realism, Stage Three) Expressiveness, Stage Four) Style and Form, and Stage Five) Autonomy. However, the question "Is it beautiful?" was not asked. Parsons (1987) states

the five stages identified in the study generally parallel Kohlberg's stages of moral development, but the study did not present quantitatively correlated results. Parsons does not address any differences between an individual's developmental reasoning about one work of art and the other works, either qualitatively or quantitatively. Parsons does not describe how conceptions of the understanding of works of art and conceptions of justice-reasoning relate to one another at each stage, nor does he calculate empirical correlations between the two domains at each stage or level. Commons (1998) and Erdynast (1974) find that the nature and stage of the task presented to subjects have a very strong influence on the stage of problem-solving that is elicited.

Erdynast, Armon and Nelson (1978) studied the structures of conceptions of the beautiful using a single lithograph by Picasso, asking the central question in aesthetic theory (Adler, 1981), "Is it beautiful?" along with the clinical interview structural reasoning probe question, "Why or why not?" A reproduction of Picasso's four line etching of buttocks from an illustration for Ovid's Metamorphosis (1931) was used to elicit judgments of whether the etching is beautiful or not. Results of the study organized the responses into cognitive-developmental levels (also called structural-developmental levels), ranging from egoistic judgments of taste to principled judgments of the beautiful.

The current study uses four paintings by Picasso and one sculpture by Michelangelo to research structural-developmental conceptions of the beautiful. The particular paintings by Picasso were selected to enable researchers to distinguish higher level conceptions of the beautiful from lower ones. Picasso's *Woman Ironing (1904)* was selected for its depiction of the state of resignation. Unlike Degas' *Woman Ironing (1884)* or *Women Ironing*

(1884) where the women ironing use their strength to do so and which are period pieces, Picasso's subject has no strength and her hands are set in opposing directions so that she cannot iron. Picasso's painting represents a woman without hope, completely devoid of energy, indifferent to how disheveled she looks. She is only able to remain standing because she holds the iron with one arm completely extended so as to prevent her collapse. Picasso depicts extreme extension of her left shoulder, even beyond Van Gogh's extended limbs and depicts her with barely the strength to preserve her upright position.

Picasso's *Weeping Woman* (1937) is depicted as not experiencing any relief from her grief through crying. It is the expression of a mother's heart-rending grief as was depicted in the woman holding her dead child in *Guernica*. The painting goes beyond the specifics of the senseless, inhuman bombing of a peaceful Basque town by Nazi planes in support of General Franco by its generality to any circumstance of a mother's suffering for her children who have been kidnapped as victims of human trafficking or sexual slavery. The painting is self-sufficient in conveying the vehemence of the loss. The intensity of the colors heightens the intensity of her distress and her unavailing grief.

Picasso's *Les Demoiselles d' Avignon* (1907) involves Cubist representations, not images of actual persons. It is a painting that involves artistic originality and the discovery of new limits in the geometric analysis and representation of female form (especially faces, torsos, breasts and noses), the mask, and includes multiple viewpoints integrated into a single viewpoint for the viewer. Picasso has examined the limits of being revealing and exposed, from minimum to maximum in the various poses adopted by the women while preserving their genital privacy. The anatomy of the

Table 1. four domains of human development

I. THE REAL	II. THE RATIONAL	III. THE REASONABLE	IV. THE BEAUTIFUL
the pursuit of knowledge, the understanding of the factual reality	conceptions of the worthwhile final ends, aims, interests, attachments to persons and associations	conceptions of justice and right that involve obligations, duties and liberties in situations of competing claims	contemplation and fashioning of beautiful objects
level 4			**level 6**
sub stage 5 original theory	**level 6**	**level 7**	cubism and picasso
sub stage 4 evaluative schools of thought	rational intuitionism	justice as fairness	**level 5**
sub stage 3 schools of thought	**level 5**	**level 6**	origins of new schools of art
sub stage 2 multiple theoretical perspective	the good of political life: the good of the free and	rational intuitionism	**level 4**
sub stage 1 formal operations theoretical perspective	**level 4**	**level 5**	classicism; artists of distinction, e.g. rembrandt, michelangelo
	complementary astute individualities	political liberalism	**level 3**
level 3	**level 3**	**level 4**	conventional art; e.g. family portraits
concrete operations	concrete values of primary social groups	the morality of policies and principles of large complex organizations	**level 2**
level 2	**level 2**	**level 3**	art of hedonistic satisfaction of monetary gain
pre-logical (intuitive)	self-interest, instrumental opportunism	the concrete morality of primary social groups	**level 1**
level 1	**level 1**	**level 2**	children drawing; simplistic, pleasant faces with attractive features
sensory motor	subsistence, support	egoistic conception of justice	
		level 1	
		a. morality of forceful and coercive power	
		b. morality of authority	

figures is represented geometrically in *Les Demoiselles d' Avignon* with the use of the triangle, curve, cone, circle, square, cylinder, rectangular plane, and a trapezoidal form. The painting is a study of five women in the reception room of a working-class brothel, four of them vying for a client. The painting involves dynamic sequences of progression of representation from the most naturalistic to the most abstract one.

Michelangelo's *Pietà* (1498-1500) sculpture at St. Peter's Basilica depicts Mary's resignation and acceptance of Christ's death. Since Mary is able to hold a body with only one arm, not even underneath the body to hold it up, it is sculpted to convey minimum mass of Christ's body. *The Blind Man's Meal* (1903) depicts an individual who is enormously strong as characterized by the largeness of his shoulders and arms, and especially by the strong well-defined muscles of his neck. He reaches with the most delicate touch possible for his wine goblet so as to not overturn it. Maximum is represented by strength that is conveyed by the neck muscle and minimum is conveyed by the absence of force and gentleness of his reach.

» RELATIONS BETWEEN DEVELOPMENT OF CONCEPTIONS OF THE BEAUTIFUL AND MORAL DEVELOPMENT

Marshall (1922) and Moore (1995) address relations between aesthetic and moral judgment, but not as empirical or cognitive-developmental studies with actual data. Two empirical studies have been conducted to examine relationships between structures of moral and aesthetic development among adult subjects, as they resolve moral dilemmas and articulate their contemplations of works of art (Erdynast, Armon & Nelson, 1978; Parsons, 1987). Both of these empirical studies identified five levels of aesthetic development related to Kohlberg's stages of justice-reasoning (Colby and Kohlberg, 1987).

Erdynast, Armon, and Nelson (1978) examined the structures of judgments of the beautiful with 50 subjects ranging in age from five to sixty-eight, and identified conceptions of the beautiful into levels ranging from egoistic judgments of taste to principled judgments of the beautiful. Insufficient but necessary conditions between structural-developmental levels of moral development (Kohlberg, 1969), and developmental levels of conceptions of the beautiful seem to be present. Those conditions require attaining certain levels of moral development, which also contain types of thinking that have effect on the construction of particular levels of conceptions of the beautiful. The Erdynast, Armon, and Nelson (1978) study found that adulthood conceptions of the beautiful were never higher than adulthood conceptions of the good (the domain of worthwhile final aims, ends, interests and attachments, i.e., the good) and the just, but were sometimes lower, though supporting quantitative data were not presented. Developmental levels of conceptions of the beautiful seem to fit Piaget's criteria for developmental structures (Piaget, 1965). The subjects were interviewed on one or more standardized justice-reasoning dilemmas to assess their justice reasoning level. Subjects were also interviewed within the domain of worthwhile final aims, ends, interests and attachments through an interview questionnaire about conceptions of the good that included questions about conceptions of good work, good friendship and good intimate relationships.

The Erdynast, Armon, and Nelson (1978) study had several limits: *1)* the presented problem, the four line etching, was insufficiently complex to elicit potentially higher level capacities in subjects, if they did indeed possess them; and *2)* plural works of art were not presented, so it was questionable whether the identified conceptions of the beautiful were specific to just that work of art, or were general conceptions applicable to numerous works of art. The problem presented to the subjects along with the question "Is it beautiful? And "why or why not" was "solved" by higher level capacities within Stage 5 principled art analysis capacities. Capacities higher than that were not elicited if subjects possessed them by that particular work of art. Picasso's four line etching of buttocks is "attractive" or "pretty" by most conventional standards. But a painting such as Picasso's *Weeping Woman*, used in the current study, which is not a "beautiful subject," can be viewed as "beautiful" in its representation and depiction of her grief and distress. Discernment of the beautiful and its distinction from the terms "attractive" and "pretty" may not be made by subjects if the representation and depiction of the subject by the painting is insufficiently complex to elicit complex conceptual analysis of the work of art. The Erdynast, Armon, and Nelson (1978) study had another major limit *3)* it used a conception of moral development that Kohlberg later revised (Kohlberg, Levine, & Hewer, 1983), and did not use the Rawlsian philosophical framework and separation of domains and sub-domains used in the current study (Erdynast and Rapgay (2009) see Figure 1 and 2 and Appendix B).

Results of the study identified distinct hierarchical adult conceptions of the beautiful and conceptions of the good. At younger ages, subjects only demonstrated attainment of the lower conceptions of the good and of the beautiful. Developmental levels tend to rise with chronological age; however, this tendency follows a necessary but insufficient condition, wherein higher stages do not necessarily increase with older age. Several distinct adulthood levels of attainment, which were not found during childhood or adolescent years, were identified. The adulthood developmental conceptions of the good and the beautiful seem to be homomorphically parallel to adulthood conceptions of justice reasoning. In other words, there seem to be five levels of conceptions of the good and five levels of conceptions of the just, along with five levels of conceptions of the beautiful that are somewhat parallel to one another, but also distinct from one another.

» THE DOMAIN OF MORAL DEVELOPMENT

The domain of justice and right, Domain III, is the domain of competing claims between individuals or between individuals and society. Illustrative examples of studies in this domain are those of Kohlberg's longitudinal and interventional studies (Kohlberg, 1969, 1973, 1984). Kohlberg's research uses hypothetical allocative justice-reasoning dilemmas focusing on competing claims between individuals to study the development of justice-reasoning. Competing claims are resolved through six hierarchical structures of justice-reasoning. At the highest levels, these structures are principles of equal liberty for all individuals and the principle of fairness for individuals. The study of what is called or titled "justice-reasoning development" by Kohlberg is partly determined by the specific meta-ethical categories involved in the presented problem

and the types of response by the subjects (Dewey & Tufts, 1932). Four sub-domains of moral development can be characterized as: *1)* conceptions of the good, *2)* judgments of justice-reasoning (obligations, liberties and duties), *3)* supererogatory acts which a person does for the sake of another's good at considerable cost or risk to the self. Illustrative supererogatory virtues are: compassion, magnanimity, and forgiveness, and *4)* judgments of moral character and worth, i.e., judgments of approbation and disapprobation of moral actions and character (Figure 1).

However, Kohlberg, the most influential and seminal researcher in the field of the development of moral judgment, initially declared his research was "moral development," but later retracted this broader title and narrowed his claim to a more constrained view of the development of justice-reasoning as the study of duties, obligations and liberties involved in allocative justice (Kohlberg, Levine & Hewer, 1983).

The domain of justice-reasoning involves questions that address moral requirements (Rawls, 1971) that are specified by issues of obligations or natural duties (Figure 2). The issues of justice are determined by fairness and fidelity. Rawls specifies the natural duties as ones that branch into two types: positive duties and negative duties (Rawls, 1971, p. 107; Figure 2). The positive natural duties involve upholding justice, mutual aid, and mutual respect. The negative natural duties are not to injure and not to harm the innocent. Colby and Kohlberg's interview protocols are largely compatible with these meta-ethical categories (1987). Dilemmas address moral requirements, and their related probe questions are constrained to these moral categories. When dilemmas and probe questions go beyond justice-reasoning they invoke philosophically more inclusive types of conceptions of right and moral development that involve the supererogatory virtues such as compassion, benevolence, and magnanimity (Figure 2).

As a consequence of revising his claim from having studied justice-reasoning rather than moral development, Kohlberg's (1987) scoring system should be understood to be constrained to measure justice-reasoning, not the development of a broader conception of right, though the standard research protocols include questions about relative moral worth and meta-ethical questions and punishment as well. The implications of the revision for the Colby and Kohlberg (1987) scoring manual is that the following categories in the table of The Elements and Norms for Classifying Content are not applicable since they do not apply to justice-reasoning types of moral judgments: the Modal Elements of blaming and retributing, and the Value Elements of upholding character, upholding self-respect, and serving human dignity and autonomy (Table 2). In Kohlberg's Heinz dilemma, levels of justice-reasoning involve competing claims and a judgment of which claim is right to select, such as stealing or not stealing (standard dilemma III), ordering the demolition expert or using a lottery to select which soldier is to go on a suicide mission (dilemma V), and providing an assisted suicide or declining to do so (dilemma IV). The content choice of alternatives is an important matter, not just the identified level of perspective being used by the interview subject (Selman, 1980). Without deciding on what the right course of action might be through a content choice, a level of perspective merely identifies the types of considerations a subject is contemplating. With the same developmental level held constant, subjects sometimes shift their content choice when asked to select a choice based on compassion rather than based on duties or obligations (Erdynast & Rapgay, 2009). The shift is accounted for by motivation emanating from the love of humankind upon which compassion is based, in contrast to the requirements of obligations, duties and liberties that are based on conceptions of justice.

worthwhile ends	**justice-reasoning**	**supererogatory virtues**
i.e. conceptions of the good	i.e. conceptions of obligations, duties, and liberties	e.g. compassion, magnanimity, forgiveness

LEVELS OF MORAL CHARACTER INVOLVE COMBINATIONS OF THE ABOVE THREE TYPES OF MORAL BEHAVIOR

Rawls, J. (1971). *A theory of justice.* Cambridge: Harvard University Press

Figure 1. levels of moral character
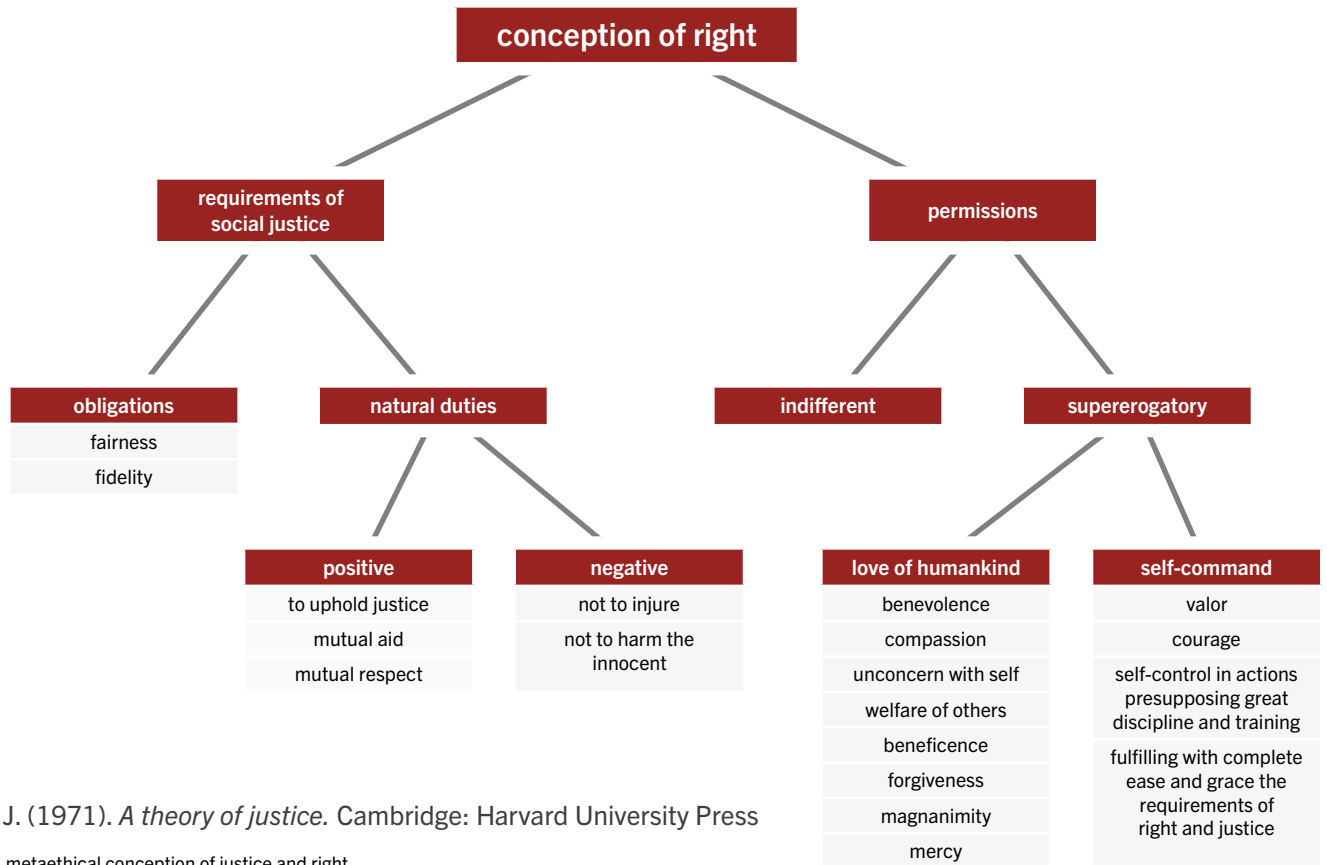
Rawls, J. (1971). *A theory of justice.* Cambridge: Harvard University Press

Figure 2. metaethical conception of justice and right

» SUBJECTS

Three groups were studied involving 180 subjects, with nearly an equal number of males and females. The age range was largely from 15 to 63. The subjects' formal education ranged from high school to doctoral level studies. Group 1 is a convenience sample of adult individuals who are returning to college to complete their Bachelor of Arts degree, taking a psychology course, and also other adult individuals who were willing to be interviewed by the students for additional units of academic credit.

Group II involves a similar population of adult students involving 19 subjects who responded to a standardized stealing dilemma (Colby & Kohlberg, 1987, Dilemma III) and a reproduction of Picasso's painting *Les Demoiselles d' Avignon*, 1907. Group II is also a convenience sample of adult individuals who are returning to college to complete their Bachelor of Arts degree, taking a psychology course, and also adult individuals who were willing to be interviewed by the students.

Table 2. the elements and norms for classifying content

the elements					the norms
modal elements	**value elements**				life *a.* preservation *b.* quality/quantity
obeying (consulting) persons or deity; should obey, get consent (should consult, persuade)	**egoistic consequences**	**utilitarian consequences**	**ideal or harmony-serving consequences**	**fairness**	property
blaming (approving). should be blamed for, disapproved (should be approved)	good reputation (bad reputation)	good individual consequences (bad individual consequences)	upholding character	balancing perspectives or role taking	truth
			upholding self-respect		affiliation
retributing (exonerating); should retribute against (should exonerate)	seeking reward (avoiding punishment)		serving social ideal or harmony	reciprocity or positive desert	(erotic love and sex)
		good group consequences (bad group consequences)	serving human dignity and autonomy	maintaining equity and procedural fairness	authority
having a right (having no right)					law
				maintaining social contract or freely agreeing	contract
having a duty (having no duty)					(civil rights)
					(religion)
					conscience
					punishment

Group III involves Picasso's contemporaries, who had published their commentaries, analyses and critiques of *Les Demoiselle d' Avignon* or expressed their reactions in interviews. A Museum of Modern Art publication on the painting Rubin, Picasso, Seckel and Cousins (1994) contains these analyses, interviews and critiques. The number of subjects for Group III is 25.

» METHOD OF SAMPLING

A convenience group of subjects was used for Group 1 and Group 2. The subjects are undergraduate students in several psychology classes taught by the primary researcher. Students were invited to participate in the studies as part of their term projects for the academic credit. For an additional unit of credit, students were invited to interview someone whom they know who was not a student as long as that individual also consented. The research objectives, and the findings do not make any claims about representative generalizations of mean, modal, average or range of levels expected to be exhibited by similar individuals. The research projects have been aimed at a discovery phase of levels of conceptions of the beautiful that are characteristic of adult development.

The convenience samples differ from preponderant majority of convenience samples with undergraduates. The age range in most of those undergraduate research samples is 17-21. The average age of the sample used in these studies is older, since the age range of the sample groups in these studies is 15 to 63 and above. Additionally, secondary subjects whom the primary subjects interviewed included 5 individuals with master's degrees and 2 Ph.D.'s, and 4 LLB degrees as well.

The 25 analyses, commentaries and critiques selected for Group III were all of the ones that had sufficiently lengthy expressions that could be scored as data that could be matched to representative data at various levels of conceptions of the beautiful using the scoring manuals used in this study.

» STUDIES

Study 1 involves research subjects' responses to whether four works of art are beautiful, using works by two artists, Picasso and Michelangelo. Reproductions of four works of art were presented: *1)* Picasso's *Blind Man's Meal, 1903*, *2) Woman Ironing, 1904*, *3) Weeping Woman, 1937*, and *4)* Michelangelo's *Pietà* at St. Peter's Basilica, 1499-1500. Group I responded to a modified standard moral judgment interview (Colby & Kohlberg, 1987; Erdynast, 1974) which involves two alternative procedures for selecting a soldier to go on a suicide mission to save the lives of the remaining Marines (Colby & Kohlberg, 1987, Dilemma V). This group involved 91 subjects who responded to reproductions of 3 Picasso paintings and a Michelangelo sculpture of a Pietà; 77 subjects responded to the modified Korean Dilemma, and 83 subjects responded to a second moral development instrument—the fidelity/infidelity questionnaire.

In order to identify general conceptions of the beautiful, subjects' responses are first examined for developmental level consistency across multiple works of art; and second, for developmental level consistency across two artists' works, specifically, three paintings by Picasso and a sculpture by Michelangelo.

Study 2 involves research subjects' responses to Picasso's painting *Les Demoiselles d' Avignon*, a painting theoretically at a level of aesthetic complexity above all the other paintings used in this study and the sculpture in Study 1 used as tasks for developmental research purposes in the domain of the beautiful. These subjects also responded to a justice-reasoning dilemma.

Study 3 involves the analysis of developmental levels of conceptions of the beautiful, inherent in the 25 published analyses of *Les Demoiselles d' Avignon* and reactions elicited in interviews of their reactions to the painting (Rubin, Seckel & Cousins, 1994).

» MORAL REASONING DILEMMAS

The study examines structural-developmental levels of: *1)* whether subjects think a painting is beautiful or not beautiful (the content-choice), and why or why not (developmental structures of conceptions of the beautiful); *2)* the relations between developmental level of conceptions of the beautiful and structural-developmental levels of moral decisions specified by a) conceptions of justice-reasoning and b) conceptions of compassion.

At least one standardized moral dilemma was administered orally or distributed on paper to subjects for their responses. These particular dilemmas are modifications of three of the Colby and Kohlberg's (1987) Standard Moral Judgment Interview (MJI) dilemmas that can elicit high levels of moral reasoning (Erdynast, 1974) (Appendix A). These MJI dilemmas were modified so that there would be a justice-reasoning component (specifically the administration of questions of obligation and duties) and a super-erogatory component (specifically compassion). These components are stringently matched to the meta-ethical categories (Appendix B) based on Rawls' theory (Rawls, 1993, 1999). Within the domain of the just and right there are: *1)* a set of requirements specified by 1a) the obligations and 1b) the positive and negative natural duties, and *2)* supererogatory excellences or virtues, such as compassion, magnanimity or forgiveness. Meta-ethical judgments, commutative justice judgments, and comparative moral worth judgments have been removed from the measure of justice-reasoning, since all of these are distinct from justice-reasoning itself and contrary to Kohlberg's own theory (Kohlberg, Levine, & Hewer, 1983).

In addition to standardized dilemmas, subjects in Group 1 were administered a fidelity/infidelity structural interview (Appendix C). Interviewing subjects on their developmental conceptions of fidelity and infidelity examines conceptions emanating from Domain II: conceptions of the good and Domain III conceptions of duties and obligations rooted in justice-reasoning, and also conceptions of compassion rooted in the morality of the love of humankind. Particular questions that seem central to issues of fidelity and infidelity are addressed in these questionnaires: *1)* Is fidelity a duty or obligation in a marriage? *2)* Can adultery emanate from a duty to self? *3)* If a spouse has engaged in adultery, is there a duty to inform the other spouse? *4)* Should the adultery be revealed if the individual is motivated by compassion? (Appendix C) Administration of questionnaires that address issues in the domains and subdomains of the good, and the just and right enable the study of these domains in their relation to the domain of the beautiful.

The moral judgment instrument for the levels of conceptions of fidelity/infidelity was devised empirically. The results of a correlational analysis indicated that levels of justice reasoning elicited from the Korean Dilemma were positively correlated with those elicited from the Fidelity/Infidelity questionnaire, $r(75) = .675$, $p<.01$. Similarly, levels of compassion reasoning were correlated between the Korean Dilemma and the Fidelity/Infidelity questionnaire with almost an equally strong positive correlation, though based on a somewhat smaller sample or respondents, $r(54) = .596$, $p<.01$.

This study presumes the capacity for several distinct developmental levels of conceptions of the good, conceptions of the justice, conceptions of compassion and conceptions of the beautiful within individuals and across different individuals. Gathering data for assessment of a subject's developmental conceptions has two interrelated components: *1)* administration of moral dilemmas designed to elicit an individual's high-level conceptions of moral thinking and conceptions of the beautiful, and *2)* probing questions designed to elicit an individual's reasons for choices of what is good, just, compassionate and beautiful (Erdynast, 1974; 1984).

The research engaged in three similar-theme studies examining conceptions of the beautiful, with works of art of increasing degrees of complexity. Unity in the research studies was based on standardized analysis of data provided by three cross-sectional groups of subjects. A repeated survey was administered two times with cross-sectional groups of subjects. Standardized questions were asked of the first two study groups using first a single art work, and then four works of art by two artists, Picasso and Michelangelo, then *Les Demoiselles d' Avignon* by Picasso.

The research study engaged in a discovery phase examination. In a parallel with Kohlberg's dissertation study which, at the discovery phase, identified stages of justice-reasoning in a cross-sectional sample of males (Kohlberg, 1958). The objectives of findings at discovery phases are to identify and characterize each of the stages or levels development. Subsequent phases of research study various representative samples of populations including, children, adolescents, adults and specialized groupings and make generalizations about the various categories of subjects.

» SCORING MANUALS

Scoring manuals were constructed in order to obtain reliability in scoring data. The scoring manuals for Picasso's *Blind Man's Meal*, *2) Woman Ironing, 3) Weeping Woman, 4)* Michelangelo's *Piet*à at St. Peter's Cathedral, and *5) Les Demoiselles d' Avignon* are organized into three types of content or sections for each level of conceptions of the beautiful. These levels include: *1)* the structural response to whether the painting is beautiful or not, *2)* the characteristic misconceptions or deforming assimilations, and *3)* the typical attributions that cannot be objectively supported within the painting.

The moral development scoring instruments utilize the scoring manuals from the Erdynast and Rapgay (2009) study of developmental conceptions of compassion, which are more encompassing than Kohlberg's Moral Judgment scoring manuals of justice-reasoning development. These scoring manuals separate justice-reasoning and compassion from *1)* meta-ethical reasoning, *2)* commutative justice punishment theory, and *3)* judgments of moral worth.

Interview responses were scored at either discrete levels such as Level 2 or Level 3, or at mixed levels such as Level 2.5, which indicates the subject used some combination of both Level 2 and Level 3. Highest manifested level and lowest manifested level were recorded for each subject without calculating a weighted mean based on frequency of each level utilized by the subject. The theoretical premise at the basis of such scoring procedures is that in any real world decision-making the subject makes a moral judgment at the highest level of moral reasoning capacity available to that subject, or alternatively the subject can revert to lower level constructions.

» RELIABILITY

Two raters independently scored conceptions of the beautiful data. The inter-judge reliability was 82% unanimous consensus across the five works of art, and 86% of responses scored within a half level or the same level as one another.

Two raters independently scored the moral judgment interview data. The inter-judge reliability is high: 86% of the interviews were assigned identical scores and 96% interviews were scored at the same level or within half a level of each other.

» RESULTS

The research study successfully: *1)* identified a highest level conception-principled basis for judgments of the beautiful that had not been identified in previous studies, *2)* characterized six developmental levels of conceptions of the beautiful, *3)* constructed scoring manuals for analysis of data and *4)* found relations between developmental levels of moral reasoning and developmental conceptions of the beautiful. The number of subjects studied is 180. Since the questionnaires were distributed for subjects to respond to, it was found that a good number of interviews were only partially answered. Among all interviews that were answered some had data that could not be scored on certain dilemmas or questions.

» I: QUALITATIVE FINDINGS

Results of the study identified the following 6 adult levels of conceptions of the beautiful:

Level 1 pre-aesthetic

Not beautiful: Beauty cannot be associated with the painting *Les Demoiselle d' Avignon*, 1907, since it cannot be comprehended by the viewer. It is viewed as strange. Individuals guess what the painting is about, when it does not make sense to them.

Level 2 egoistic tastes

Beautiful: The beautiful is equated with one's personal tastes. Individuals either like a painting or they do not like it.

Not beautiful: Absence of beauty derives from not liking the shapes, the looks of the faces or an overall global feeling about a painting.

Level 3 cultural views of the beautiful

Beautiful: The beautiful is equated with attractiveness and positive emotions such as depictions of happiness.

Not beautiful: Absence of beauty derives from qualities of sadness and unhappiness. Disturbing aspects include ugliness, strange angularity, disproportion and dissonance.

Level 4 formal art analysis

Beautiful: There is comprehension of the subject matter but without comprehension of Cubist representation of geometric form. The beautiful incorporates formal aesthetic properties such as: order, definition, perfection, brightness unity among the disparate parts, symmetry, integrity, and due proportion or harmony.

Level 5 principled art analysis

Beautiful: Comprehension of the aesthetic principles utilized in major artistic works. The critically evaluation of various works of art, their excellences, their limitations and their weaknesses in relation to other artists and other artistic styles across major schools of artistic thought.

Table 3. range of variability of conceptions of the beautiful (highest - lowest

range of variability	frequency	percentage
0	75	50.7%
0.5	50	33.8%
1	14	9.5%
1.5	9	6.1%
total	148	100%

triangle, curve, cone, circle, square, cylinder, rectangular plane, and a trapezoidal form. The painting is a study of five women in the reception room of a working-class brothel, four of them vying for a client. The painting involves dynamic sequences of progression of representation from the most naturalistic to the most abstract one. The figure on the lower right, illustrates with completely Cubist geometric multiple perspectives best, depicted simultaneously with a three-quarter back view from the left, another from the right, and a frontal view of the head.

The mask in *Les Demoiselles d' Avignon*

Women working in a brothel are not indifferent to who is chosen. The women have good reason to be wearing masks. The *Demoiselles* either wear actual masks or masked expressions--faces that are fixed, blank and expressionless. Picasso had an insight about the mask; that the mask is used historically, as a fetish to ward off the unknown—"the evil spirits". The painting presents a timeline of

Table 4. *Michelangelo's Pieta* levels and content choice

content choice	conceptions of beautiful exhibited					
	2	percentage	3	percentage	4	percentage
yes	3	37.50%	38	80.85%	63	91.30%
no	5	62.50%	8	17.02%	5	7.25%
undecided	0	0.00%	1	2.13%	1	1.45%
total	8	100.00%	47	100.00%	69	100.00%

Table 6. *Blind Man's Meal* levels and content choice

content choice	conceptions of beautiful exhibited					
	2	percentage	3	percentage	4	percentage
yes	0	0.00%	16	41.03%	45	77.59%
no	13	100.00%	22	56.41%	11	18.97%
undecided	0	0.00%	1	2.56%	2	3.45%
total	13	100.00%	39	100.00%	58	100.00%

Level 6 content-dependent principled conceptions of the beautiful, e.g., Cubism

Beautiful: The famously known painting *Les Demoiselles d' Avignon* should be understood with a different title, the one Picasso gave to it (*My Brothel*) rather than the title by which it is universally known. *Les Demoiselles d' Avignon, Avignon* involves Cubist representations, not images of actual persons. It is a painting that involves artistic originality and the discovery of new limits in the geometric analysis and representation of female form (especially faces, torsos, extended elbows, breasts and noses), the mask, and includes multiple viewpoints integrated into a single viewpoint for the viewer. Picasso has examined the limits of being revealing and exposed, from minimum to maximum in the various poses adopted by the women. The anatomy of the figures is represented geometrically in *Les Demoiselles d' Avignon* with the use of the

the history of the mask from ancient times through early history to the present. The ancient Africanized mask on the figure in the top right characterizes a fierce baboon and is used as a fetish to ward off dangers. A mask may help to preserve self-respect and self-esteem and help avert descending into cynicism and despair. The mask on the figure the most to the left, in an Egyptian type stance, is Carthaginian. The Cubist mask originated with this painting. No one had previously seen such a mask because it was just invented by Picasso. It is the very first Cubist mask ever. The figure in the top right is a design, paralleling Matisse's Blue Nude, wearing an African type derived mask. But tribal art is not based on content-dependent Cubist conceptual principles. Cubist noses are noses seen from two perspectives and presented in a single one. The noses are presented with simultaneous lateral and frontal views.

Table 5. *Woman Ironing* levels and content choice

content choice	conceptions of beautiful exhibited					
	2	percentage	3	percentage	4	percentage
yes	1	11.11%	26	57.78%	56	83.58%
no	8	88.89%	18	40.00%	9	13.43%
undecided	0	0.00%	1	2.22%	2	2.99%
total	9	100.00%	45	100.00%	67	100.00%

Table 7. *Weeping Woman* levels and content choice

content choice	conceptions of beautiful exhibited					
	2	percentage	3	percentage	4	percentage
yes	0	0.00%	10	26.32%	40	59.70%
no	20	95.24%	26	68.42%	24	35.82%
undecided	1	4.76%	2	5.26%	3	4.48%
total	21	100.00%	38	100.00%	67	100.00%

Table 8. *Les Demoiselles d'Avignon* levels and content choice

content choice	conceptions of beautiful exhibited							
	2	%	3	%	4	%	5	%
yes	0	0.0%	3	50.0%	17	58.6%	17	73.9%
no	2	100.0%	3	50.0%	11	37.9%	6	26.1%
undecided	0	0.0%	0	0.0%	1	3.5%	0	0.0%
total	2	100.0%	6	100.0%	29	100.0%	23	100.0%

Table 10. education and levels of conceptions of the beautiful exhibited

		levels of conceptions of the beautiful					
		2	2.5	3	3.5	4	total
education	high school	1	1	12	6	1	21
	college	5	2	24	69	17	117
	masters	0	0	2	2	1	5
	ph.d.	0	0	0	0	2	2
	total	6	3	38	77	21	145

» II. QUANTITATIVE FINDINGS

Structured wholes

Different individuals in Study 1 responded to the four works of art as beautiful or not, with developmental level consistency across various works of art by the two different artists. These structural-developmental responses of conceptions of the beautiful have hierarchical characteristics. For example, subjects who used level 2 conceptions of the beautiful on one art work, tended to use the same level on the others, or a proximate level of 2.5, which is a mixed combination of level 2 and level 3, or a level 3 at the most. When subjects possessed level 4 conceptions of the beautiful, they tended to use that developmental level and proximate levels to express their appreciation of all, or most of the four works of art. Appreciating a work of art as beautiful is not merely a situation-specific judgment about a particular style or artist, or the content of the subject of the work of art, such as a beautiful subject.

Range of exhibited variability

Task complexity is evident in the level elicited for "no" responses to the paintings and for "yes" responses to the paintings as well. Where the response is "not beautiful" or "ugly," there are a few instances of developmentally lower reasons for that aesthetic judgment, due to inability to comprehend the level of the task complexity of the aesthetic issues addressed in the complex paintings. For those who comprehend or understand the painting there is a monotonically increasing percentage of individuals who respond that the painting is beautiful with each increasing level.

Education and conceptions of the beautiful

Usable data from subjects were examined. The relationship between subjects' level of formal education and their conceptions of the beautiful were examined. There is a significant association between education and hierarchical developmental conceptions of the beautiful, $p=.007$. As individuals' level of formal education level rises, their developmental level of conceptions of the beautiful tend to rise to higher levels as well, but higher levels of formal education do not necessarily result in higher levels of attainment of conceptions of the beautiful.

Table 9. education

	freq.	%
high school	23	14.5
college	129	81.1
masters	5	3.1
Ph.D.	2	1.3
total	159	100
missing	21	
total	180	

Developmental of conceptions of the beautiful and moral reasoning

Relations between highest level conceptions of the beautiful exhibited, conceptions of the good, and conceptions of justice-reasoning were examined for 144 subjects. 83.4% of the subjects exhibited the same level or 0.5 level variation of development for both moral development and conceptions of the beautiful.

Table 11. conceptions of the good minus conceptions of the beautiful

range of variability	freq.	%
−1.0	1	1.5
−0.5	20	30.3
0.0	22	33.3
0.5	13	19.7
1.0	9	13.6
2.0	1	1.5
total	66	100.0

Table 12. conceptions of justice-reasoning - conceptions of the beautiful

range of variability	freq.	%
-1.0	1	0.9
-0.5	27	25.5
0.0	52	49.1
0.5	17	16.0
1.0	8	7.5
2.0	1	0.9
total	106	100.0

Justice-reasoning and conceptions of the beautiful

Results of a paired samples ttest ($t(91) = 1.472$, $p=.145$) indicate that the mean developmental level for the conceptions of justice-reasoning is not significantly greater than the mean developmental level of conceptions of the beautiful.

Relations between conceptions of justice-reasoning and the conceptions of the beautiful

Relations between the conceptions of justice-reasoning and conceptions of the beautiful were examined. As the level of justice-reasoning development rises, there is a tendency for the level of conceptions of the beautiful to also rise.

Highest conceptions of the beautiful

Relations between highest level of conceptions of justice-reasoning and mean of highest level of conceptions of the beautiful exhibited by subjects were examined. As the level of justice-reasoning rises, there is a tendency for the mean of level of conceptions of the beautiful to also rise.

Table 13. paired samples statistics

	M	N	SD	SE
highest conceptions of justice-reasoning	3.359	92	0.401	0.042
highest conceptions of the beautiful	3.288	92	0.493	0.051

Table 14. conceptions of justice-reasoning versus conceptions of the beautiful

		highest levels of conceptions of the beautiful						
		2	2.5	3	3.5	4	5	total
highest conceptions of justice-reasoning	2.5	1	1	0	0	0	0	2
	3	4	1	19	15	1	0	40
	3.5	0	1	6	23	2	0	32
	4	1	0	2	6	9	0	18
	4.5	0	0	0	0	2	10	12
	5.5	0	0	0	0	0	1	1
	6	0	0	0	0	0	1	1
	total	6	3	27	44	14	12	106

» DISCUSSION

The domain of conceptions of the beautiful

The finding of general consistency of the use of developmental structures with the same or proximate levels of conceptions of the beautiful across multiple works of art and across two artists is of considerable import. Finding consistency across multiple works of art was not achieved in the Erdynast, Armon, and Nelson (1978) study since only a single work of art was then used to elicit structural-developmental conceptions of the beautiful. Conceptions of the beautiful could therefore conceivably have been those of work of art-specific or artist-specific responses.

The finding of consistency in the first two studies supports the premise that levels of conceptions of the beautiful are not just situation specific judgments of the beautiful, but are, rather, general structures of thought and feeling that are invoked in generating all judgments of the beautiful. The structures thus meet the criterion of generality. The study also finds that 73 of the 148 subjects, 49%, exhibit multiple hierarchical levels of conceptions of the beautiful. Most of the subjects, 94%, exhibit proximal developmental level consistency, plus one or minus one on other works experienced as beautiful and 51% of 148 subjects use their highest level of conception of the beautiful across various works of art, paintings or sculpture by more than one creative artist. In 85% of instances, the difference with the predominant modal level is only .5. In only 10% of instances do subjects manifest 1.5 differences with the predominant modal levels. These findings imply that the developmental levels of conceptions of the beautiful are not a set of mere typologies where individuals switch between types of judgment without a particular order, but, rather constitute and ordered and directional hierarchical sequence of invariant structures.

Relations between domain of conceptions of the justice, domain of the conceptions of the good and the domain of conceptions of the beautiful

As the level of moral development rises, there is a tendency for the levels of conceptions of the beautiful also to rise. Increasing age is a necessary but insufficient condition for attainment of the higher levels of conceptions of the beautiful. Misconceptions about the subject matter of paintings and sculpture are not limited to only children—they are prevalent in adults as well. Adulthood levels of conceptions of the beautiful are plural—there is not just one singular adulthood level.

The adulthood conceptions of the beautiful seem to be homomorphic parallel to adulthood developmental conceptions of the good, and conceptions of the just and the right. The developmental conceptions of the beautiful are distinct and separate from the moral structures, but hierarchically rise in parallel fashion to the moral structures. In 49% of instances, the level of conceptions of the beautiful was the same as the level of conceptions of justice. In 91% of instances, there were plus or minus a half level correspondence between the level of conceptions of the beautiful and the level of conceptions of justice. In 33% of instances, the levels of conceptions of the beautiful are identical to levels of conception of the good. In 83% of instances, the levels of conceptions of the good are within a half level, plus or minus, to levels of conceptions of the beautiful. And, in 98% of the cases, the levels of conceptions of the good and the levels of conceptions of the beautiful are within one level, plus or minus, to one another. There seem to be five levels of conceptions of the good, and five levels of conceptions of the just, along with five levels of conceptions of the beautiful, and uniformity of general development across the domains, though not completely at just one level. Data on level 6 justice structures and level 6 conceptions of the beautiful are too sparse to make generalizations about their relationships.

Table 15. conceptions of justice-reasoning and conceptions of the beautiful

	levels of conceptions exhibited				
highest conceptions of justice-reasoning	2.5	3	3.5	4	4.5
mean of conceptions of the beautiful	2.25	3.10	3.41	3.61	4.83

Asking if a painting like Picasso's *Les Demoiselles d' Avignon* is beautiful is different from asking 'is it right?' or 'is it real?' or 'is it good?' Each of those four questions represents a different fundamental question in a distinct domain of its own. What is right is not reducible into fewer than two questions, which are *1)* what is good and *2)* what is just and right within social contract

RELATIONS BETWEEN CONCEPTIONS OF JUSTICE-REASONING AND CONCEPTIONS OF THE BEAUTIFUL
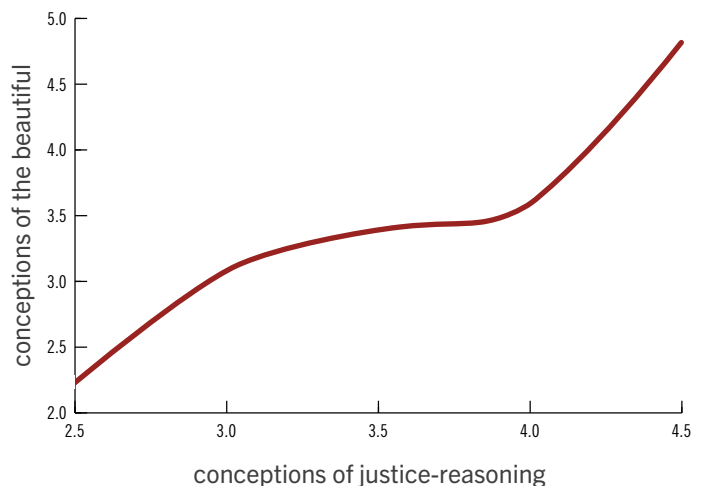
Figure 3. relations between conceptions of justice-reasoning and conceptions of the beautiful

theory, but can be reducible to what is good within utilitarian theory (Marshall, 1922; Rawls, 1971). Utilitarian moral theories permit a three-domain model of the real, the right, and the beautiful, since conceptions of the right are ones that are secondary to conceptions of the good and are defined as those that maximize the good (Rawls, 1971). Utilitarianism is teleological philosophical moral theory. Within utilitarianism the good is primary, and the right consists of maximizing the good. Utilitarianism also does not have the supererogatory virtues, such as compassion, benevolence, magnanimity or forgiveness, within moral conceptions of the right. Utilitarian theories incorporate the good and the right together rather than making distinctions between them (Marshall, 1922; Rawls, 1971). The four distinct questions cannot be conflated with one another within social contract theories, which separate the good from the right and constrain or forbid the good if it violates the right (Rawls, 1993).

In a Kantian framework, the beautiful is a domain and category of judgment of its own, *sui generis*, one that is irreducible to other types of judgments for analyzing artistic beauty (Kant & Meredith, 1911). Kant's principle of subjective universality incorporates criteria for a normative end point of objective judgments of the beautiful. The domain of the beautiful is what Fry (1966) refers to as Kantian autonomy in aesthetics.

The types of developmental levels of "hard" structural stages central to the focus of Piaget's (Inhelder & Piaget, 1966) and Kohlberg's studies (1969) are ones whose qualities are structural and progress through invariant sequences of construction that transform previous structures rather than qualitatively add to them. Progression cannot occur, for example, from level 1 to level 4 directly. There is no skipping of sequential levels. Each stage provides a platform from which the next levels are constructed. Individuals can, however, revert from higher level capacities to lower levels of functioning. The four criteria for hard structural stages (Kohlberg & Armon, 1984) are all applicable to the four domains. Developmental stages or levels incorporate qualitative changes in thought and affect that are constructed by the individual. Each stage forms a generalized structured whole that is applied in solving all similar issues. These qualitative changes in thought and affect occur in an invariant sequence and transform the previous highest developmental structure into a new hierarchically higher structural whole. While the research study was a cross-sectional one, the connection between the findings in the domain of the beautiful to ones in the subdomain of justice-reasoning suggests that longitudinal studies would confirm ontogenetic progression through the six structural developmental conceptions of the beautiful. ■

REFERENCES

Adler, M. J. (1981). *Six great ideas: Truth, goodness, beauty, liberty, equality, justice: ideas we judge by ideas we act on*. New York: Macmillan.

Baldwin, J. M. (1915). *Genetic theory of reality*. New York: G. P. Putnam's Sons.

Colby, A., & Kohlberg, L. (1987). *The measurement of moral judgment*, Vol.1: Theoretical foundations and research validation. New York: Cambridge University Press.

Collingwood, R. G. (1958). *The principals of art.* London: Oxford University Press.

Commons, M. (1998). Hierarchical Complexity of Tasks Shows the Existence of Developmental Stages. *Developmental Review, 18* (3), 237-278.

Commons, M. L. & Richards, F. (2003). Four Postformal Stages, in The Handbook of Adult Development, the Springer Series in Adult Development and Aging, Demick, J. and Andreoletti, C. eds.

Danto, A. (1981). *Transfiguration of the commonplace: A philosophy of art.* Cambridge, Mass.: Harvard University Press.

Dewey, J. (1980). *Art as experience*. New York: Perigee Books.

Dewey, J. and Tufts, J. H. (1932). *Ethics*. New York: Holt.

Dor., S. R., & Musée Picasso (Antibes, France). (1960). *Picasso in Antibes*. London: Lund, Humphries.

Erdynast, A. (1974). *Improving the adequacy of moral reasoning: an exploratory study with executives and doctoral philosophy students*. (Unpublished doctoral dissertation). Harvard University, Cambridge.

Erdynast, A. (1984). A Rawlsian view of Kohlberg's conception of stage 6 justice-reasoning. In M. Commons, F. Richards, & Armon (Eds.). *Beyond formal operations, Volume 2: Late adolescent and adult cognitive development*. New York: Praeger.

Erdynast, A., Armon, C., & Nelson, J. (1978). Cognitive developmental conceptions of the true, the good, and the beautiful. *The eighth annual proceedings of Piaget and the helping professions.* Los Angeles: University of Southern California Press.

Erdynast, A., & Rapgay, L. (2009). Developmental levels of conceptions of compassion in the ethical decision-making of Western Buddhist practitioners. *Journal of Adult Development, 16* (1), 1-12.

Erikson, E. H. (1959). Growth and crises of the healthy personality. *Psychological Issues*, *1* (1), 50-100.

Fry, E. (1966). *Cubism.* London: Thames & Hudson; New York: McGraw-Hill.

Gardner, H. Winner, E. & Kircher, M. (1975). Children's Conception of the Arts. *Journal of Aesthetic Education*, (9)3, 60-77.

Giri, S., Harrigan, W., & Commons, M. (2014). *There is only one stage domain*. In Press.

Inhelder, B., & Piaget, J. (1966). *The growth of logical thinking from childhood to adolescence: An essay on the construction of formal operational structures*. London.

Kant, I., & Meredith, J. C. (1911). *Kant's Critique of aesthetic judgement*. Oxford: Clarendon Press.

Kegan, R. G. (1979). The Evolving Self: A Process Conception for Ego Psychology. *The Counseling Psychologist, 8* (2), 5-34.

Kohlberg, L. (1958). *The development of modes of thinking and choices in years 10 to 16.* (Unpublished doctoral dissertation). University of Chicago.

Kohlberg, L. (1969). Stage and sequence: The cognitive developmental approach to socialization. In Goslin, D. A., Handbook of socialization theory and research. Chicago: Rand McNally.

Kohlberg, L. (1973). *The claim to moral adequacy of a highest stage of moral judgment*. New York: Journal of philosophy.

Kohlberg, L. (1984). *Essays on moral development: Volume ii. The psychology of moral development*. San Francisco: Harper & Row.

Kohlberg L. & Armon, C. (1984). *Three models for the study of adult development*. In M. Commons, F. Richards & C. Armon (Eds.), Beyond formal operations: Late adolescent and adult cognitive development (pp. 383-394). New York: Praeger.

Kohlberg, L., Levine, C., & Hewer, A. (1983). *Moral stages: A current formulation and a response to critics*. Basel: Karger.

Langer, S. (1953). *Feeling and form.* New York: Charles Scribner's Sons.

Levinson, D.J. (1978). *The Seasons of a man's life*. New York: Knopf.

Loevinger, J., & Blasi, A. (1976). *Ego development: [conceptions and theories]*. San Francisco: Jossey-Bass Publishers.

Marshall, H.R. (1922). The True, the good and the beautiful. *The Philosophical Review, 31* (5), 449-470.

Moore, R. (1995). The aesthetic and the moral. *The Journal of Aesthetic Education*, 29 (2), 18-29.

Ovid, ., Allcroft, A.H., & Stout, J.F. (1931). *Metamorphoses: Book i*. London.

Parsons, M. (1987). *How We Understand Art*. Cambridge University Press, New York.

Perry, W.G. (1968). Forms of intellectual and ethical development in the college years: A scheme. New York: Holt, Rinehart & Winston.

Piaget, J. (1965). *The moral judgment of the child*. New York: The Free Press.

Piaget, J. (1965). *The construction of reality in the child*. New York: Balantine.

Rawls, J. (1971). *A Theory of Justice*. Cambridge: Harvard University Press.

Rawls, J. (1993). *Political Liberalism*. New York: Columbia University Press.

Rawls, J. (1999). *Collected Papers*. Cambridge, Mass: Harvard University Press.

Rubin, W., Seckel-Klein, H., Cousins, J., & Museum of Modern Art (1994) *Les Demoiselles d'Avignon*. New York: Museum of Modern Art.

Selman, R. (1980). *The Growth of interpersonal understanding: developmental and clinical studies*. New York: Academic Press.

Snarey, J., Kohlberg, L. & Noam, G. (1982). *Ego Development and Education: A Structural Perspective*, in Kohlberg, L. & DeVries, R. (Eds.), Child Psychology and Childhood Education: A Structural Developmental View. New York: Longman.

APPENDIX A

» THREE MORAL DILEMMAS:

Dilemma III—The Heinz dilemma

A man and his wife have recently migrated from the high mountains. They started a farm, but there was no rain and no crops grew. They didn't have enough food. The wife became sick from having so little food and could only sleep. Finally, she was close to dying.

The husband could not get any work, and the wife was too ill to be moved to another town. There was only one grocery store in the village, and the storekeeper charged very high prices because there was nowhere else for people to buy food. The husband asked the storekeeper for some food for his wife and said he would pay for it later. The storekeeper said, "No, I won't give you any food unless you pay first."

The husband went to all the people in the village to ask for food, but no one had any to spare. Desperate, he broke into the store to steal food for his wife.

Illustrative structural probe questions:

.301 Is it a husband's duty to steal the food for his wife if he can get it no other way and why or why not?

.302 Suppose the person dying is not his wife, but a stranger. Would the husband have a duty or an obligation to steal the food for the stranger, and why or why not?

.302B (optional) Suppose the person dying is his spiritual mentor (his root teacher). Would the husband have a duty or an obligation to steal the food for the stranger, and why or why not?

.303 Should he steal for the stranger if he is motivated by compassion and why or why not?

.304 Would that be the right decision and why?

Dilemma V—The Korean dilemma

During the Korean War, a company of ten marines is out-numbered and retreating from the enemy. The company has crossed a bridge over a river, but the enemy is still on the other side. Someone needs to destroy the bridge. If someone goes back to the bridge and blows it up, that person will probably not escape alive. Only the captain knows how to lead a successful retreat. The captain asks for a volunteer, but no one offers to go. If no one goes back, it is virtually certain that all will die.

The captain decides that he has two alternatives. The first is to order the demolition expert to return to the bridge and blow it up. If this man is sent back, the probability that the mission will be accomplished successfully is .8. The second alternative is to select someone to go back by drawing a name out of a hat. If anyone other than the demolition expert is sent back, the probability that the mission will be accomplished successfully is .7.

Illustrative structural probe questions:

.501 Which of the two alternatives is right and why?

.502 Which alternative should the captain select if he is motivated by compassion and why?

If ordering the demolition man is chosen:

.503 Does the demolition man have a duty or an obligation to go and why?

If the lottery is chosen:

.504 Does the person selected by the lottery have a duty or an obligation to go and why?

.505 Suppose the demolition man objects to being ordered to go on the mission, saying that he was drafted into the army and that while he chose to do riskier demolition work, he did not volunteer to go on suicide missions. He makes the claim that he has a right to have his life saved too, and while he objects to being the one ordered to save everyone else's life, he wouldn't object to going if his name was also included in the lottery and it was his name which was selected. How would these considerations affect what the captain should do, and why?

APPENDIX B

» META-ETHICAL CATEGORIES OF CONCEPTIONS OF THE GOOD, THE JUST AND RIGHT FOR INDIVIDUALS

A. Independence

1. A determinate scheme of final ends that one wants to realize for one's own sake.

2. Attachments to other persons, and loyalties to various groups and associations; affections and devotions that enable the flourishing of these sentiments.

3. A view of oneself in relation to the world in which one's religious, moral or philosophical conceptions of the good are valued and the significance of one's ends and attachments are understood.

B. Self-Authenticating source of moral claims

1. One's claims carry weight on their own without being derived from prior duties or obligations owed to society or to other persons.

2. One's claims carry weight on their own without being derived from, or assigned to, their particular social role.

3. Claims founded on duties to self.

C. Responsibility for ends

1. Individuals take responsibility for the self and the consequences of decisions and actions. Responsibility for consequences includes the willing endurance of hardship and suffering; but it must be presently acceptable in view of the expected or achieved good.

2. Individuals are capable of adjusting their aims and ambitions in the light of what they can reasonably expect and of restricting their claims in matters of justices to certain kinds of things.

3. Individuals regulate and revise their ends and preferences in light of their expectations of primary goods.

D. Requirements: The choice of regulative principles of justice that determine the reasonable terms to resolve competing moral claims
   D1. Natural Duties
     A. Positive natural duties
       1. To uphold justice
       2. Mutual aid
       3. Mutual respect
     B. Negative natural duties
       1. Not to injure
       2. Not to harm the innocent

   D2. Obligations
     1. Fairness
     2. Fidelity
E. Supererogatory Permissions
   A. The morality of self-command, which is the morality of the hero, in its simplest form, is supererogatory when the individual displays its characteristic virtues:
     1. Courage
     2. Valor
     3. Self-control in actions presupposing great discipline and training require great discipline and training.
     4. Fulfilling with complete ease and grace the requirements of justice and right
  B. The morality of the love of humankind, which is the morality of the saint, shows itself in advancing the common good in ways that go well beyond our natural duties and obligations.
     1. Benevolence
     2. Compassion
     3. A proper humility and unconcern with self
     4. Magnanimity
     5. Mercy
     6. Beneficence
     7. Forgiveness
     Appendix C
     DILEMMA VII - FIDELITY/INFIDELITY
   701. If both spouses permit sexual non-exclusivity in their marriage, do they have a
    right to do so? YES_____NO_____
   701a. Why or Why not?
   702. Is fidelity a duty or obligation in a marriage? YES_____
NO_____
   702a.Why or why not?
   703. What do you mean by fidelity?
   704. Is there a duty to reveal adultery to the other spouse?
YES_____NO_____
   704a. Why or why not?
   705. Is adultery ever right? YES_____NO_____
   705a. Why or why not?
   706. Should a spouse who has engaged in adultery be forgiven?
YES_____NO_____
   706a. Why or why not?
   707. Can adultery ever be "good", as distinct from" right or wrong"? YES_____NO_____
   707a. Why or why not?
   708. Can engaging in adultery ever emanate from a duty to self?
YES_____NO_____
   708a. Why would an individual have such a duty?
   709. What is a worthwhile sexual relationship? Why?

Deconstruction toward reconstruction: A constructive-developmental consideration of deconstructive necessities in transitions

Samuel Albertson

Harvard University

The intention of this viewpoint paper is to explore the terrain of developmental transitions, more specifically the necessity of the process of questioning or, deconstructing one's whole frame of knowing before a new frame of knowing can emerge a "reconstruction". Leaning on constructive-developmental theory, this paper seeks to define the deconstructive pattern that emerges and reemerges during developmental or "stage" transition, and shows how it is necessary to incorporate this deconstructed "stage" into a more complex system of knowing or "stage". The second portion of the paper then outlines a current example of how an epistemology can have roots in logical coherency, then become disruptive or deconstructive, then re-constructive, in the postmodern theory of social science methods commonly referred to as Critical Discursive Psychology (CDP). This theory is argued to be emerging as fifth order as defined by Robert Kegan (2010) in that it is reconstructive and not just deconstructive or, antimodernist as seen in not denying, but utilizing process, the disunified self, subjectivity, and theory re-production, as it is made clear in the argumentation of the second portion of the paper. The paper concludes in a clear affirmation of the process of differentiation and reintegration as integral for stage transition and growth not just in individual human development, but also in the social sciences.

KEYWORDS: constructive-development, stage transitions, deconstructing stages, complexity in meaning-making, critical discursive psychology

NTERDISCIPLINARY PERSPECTIVES on the growth of the human agent in biology, philosophy or epistemology, theology, psychology concur that there is a kind of rhythm to the life path, as seen through a process of loss, then expansion. Biological narratives claim that the perpetuation of living organisms occurs in and through the process of dividing and joining. Cells divide into their own entities, allowing for continual construction and repair of an organism (Maton et al., 1997). In other words, *wholeness is dependent upon periods of fragmentation*. Operating within a different narrative of theology, Greek Orthodox theologian John Zizioulas is well known for his treatise on the communitarian constitution of trinitarian anthropology (1993). He says that death is the "'natural' development of the biological hypostasis (ontological personhood), the cessation of 'space' and 'time' to other individual hypostases, the sealing of the hypostasis as individuality" (p. 51). Yet this constitutional makeup of the individuated (and individuating) human is changed only through "a kind of new birth for man," which he articulates

> *"Deconstruction insists not that truth is illusory but that it is institutional."*
>
> ⌐ Terry Eagleton (1987, ch. 6)

as salvation through baptism, believed to be a complete internal (socio-emotional and cognitive) regeneration as symbolized in the act of washing in water. The capacity for integration (communion), and thus healthy growth, comes through a fellowship with a wholly or holy Other Being. From the constructive-developmental field, Robert Kegan (1982) also references the biologically adaptive necessity of disintegration as proceeding re-integration: "biologists always tell us; adaptation is a matter of differentiation and integration" (p. 5). The ongoing rhythms of actively struggling to make, have, protect, and enhance meaning necessitates the moments of labor and loss of meaning, and indeed a loss of "the 'self' along the way" (p. 12).

These are helpful metaphors for the process of human growth. This process of individuation, of loss, of deconstruction, is a necessary step toward reintegration at more complex levels of being. Through Kegan's framework mentioned above, the meaning-organizing self encounters several moments of instability throughout the life quest. Change is a dangerous enterprise, for it entails balancing a tension between self-preservation and self-transformation (Parks, 1986). This process of psychological reorganization is, in Michael Basseches' terms, "messy" (1989, p. 198). Using a constructive-developmental perspective, I will define this deconstructive

pattern that emerges and reemerges throughout all of human life, and show how it is necessary to incorporate dissonance into a more complex harmony. In the second portion of this paper, I will then offer a current, working example of how this harmonious-to-disharmonious-to-harmonious epistemology appears in a postmodern theory of social science methods commonly referred to as critical discursive psychology. The point of this exploration is to highlight the often-overlooked theme of deconstruction and reconstruction in any growth process, looking specifically at the growth of individuals and the maturation of systems of thought or human organization.

» DISCLAIMERS

A few disclaimers are in order, however, before I engage fully in this exploration. First of all, when addressing themes of growth and change, there arises in us a tendency to overly romanticize a process of disequilibrium. This is likely due to the textual or cognitive proximity of disequilibrium to (re)equilibration, the regaining of stability, or simply change in general. When there is a disruption in balance, we cannot always assume that there will be a return to harmony. Many people across time and space are (have been, will be) in a constant state of developmental unrest, often characterized as mental illness (Kegan, 1982). Of course, the constructive promise of regeneration reminds us that instability, although necessary, need not be all-consuming; however, when addressing developmental instability in this paper, I want to avoid the use of platitudes or an over-simplification of the real struggle that can go into one's search for cohesion. Transition is often characterized by social chaos, cognitive static, or emotional vulnerability (Lewis, 1995). I also wish to clarify that this disruption in the cognitive categories (schemas) and structural wholeness of the person is a philosophical concept and does not necessarily entail a physical phenomenon (Basseches, 1989). A person who utters the oft-satirized phrase in American life, "I don't know who I am anymore," will say that disharmony is indeed a disorienting experience; but it is not because some illness has entered, or because a piece of him or her has been lost. We cannot downplay the fact that it may very well feel like this to the person in turmoil or, more accurately, the person-in-process. This "loss of the self" (Kegan, 1982; 1994) is nonetheless real in some very significant ways and will be explored below.

» MICRO AND MACRO-DISRUPTIONS IN STAGE DEVELOPMENT

This concept of a disruption in cognitive stability comes most clearly from Jean Piaget—in fact, as does the whole constructive-developmental program (Ginsburg & Opper, 1988). Piaget referred to this phenomenon as a disturbance, perturbation, or conflict (1985). His research on children's epistemologies in an educational context, as opposed to a therapeutic one, was often most focused not necessarily on their right answers, but the "wrong" answers given ever so confidently (Duckworth, 2006; Parks, 1986). Robert Selman (1980) terms this unsettlement as "conflicts, inconsistencies, or inadequacies in his or her internal system or beliefs or values"

(p. 81). This disturbance serves as an attempt at regaining composure, a re-balance in dealing with the world, a re-structuring of self in relationship with reality. This process, therefore, necessarily precedes the birth of new *forms* of knowing or a new equilibrium (Albertson, 2009). Re-equilibrium can be seen as a renewed or as Robert Kegan emphasizes, a newly-conceived, or *re*-cognized awareness of contradictions in thought (Montanegro, 1985). But before this sets in, the person may feel lost or without a formulated or identified voice. It is here that Lahey, et. al. (1988) identify the "growing edge"[1] (p. 37).

But we must not confuse the general similarities of conflicts or inconsistencies in thought with the actual *process of deconstructing* a way of knowing. The discovery of one's own cognitive limits, from ongoing or significant conflicts as termed by Selman above, leads to a process of *questioning entirely the meaning system itself*. Is there such a thing as make believe, wonder, or mystery? Is there such a thing as an authoritative voice? Is there such a thing as truth or consistency? If development is to be described as a helix of evolutionary truces (Piaget, 1974 and Kegan, 1982; see Figure 1 below), then deconstruction might be on the "down beat" of equilibrating movements or right in between plateaus of stability.
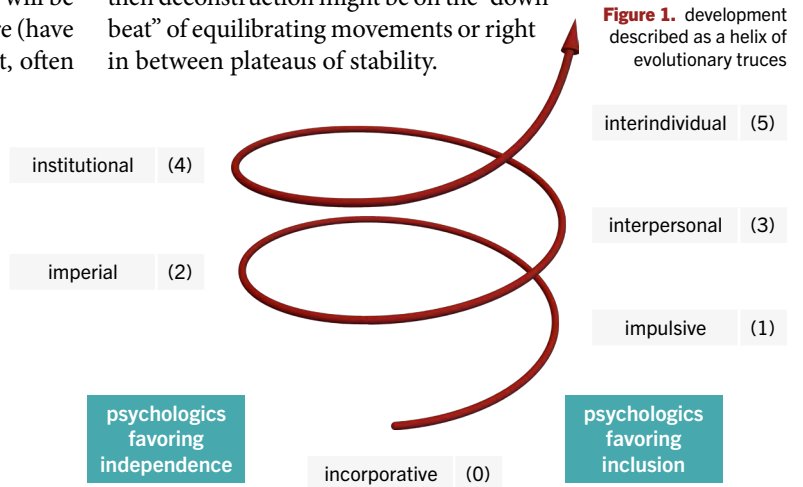
Figure 1. development described as a helix of evolutionary truces

interindividual (5)

institutional (4)

interpersonal (3)

imperial (2)

impulsive (1)

psychologics favoring independence

psychologics favoring inclusion

incorporative (0)

This deconstruction often appears in the form of invalidating the previous way of knowing. A helpful illustration of this phenomenon is that when a person who has "newly matured" into a new way of knowing, will often react strongly to or disassociate from those around them who represent that previous way of knowing. A barely-post-third order person will not want to associate with firmly-rooted third order individuals because the person will speak out at these ways of knowing which serve as representations of the person's struggled-for, newly-accomplished, and more complex frame of mind ("I can't believe *she* believes *that*!" "That person is so immature!"). This epistemology is not quite in a new order because the former way of knowing has not yet been synthesized into that still forming system. We can imagine a fourth order individual speaking about a third-order individual in a very self-authored way ("He sees the world in that way, and it is fine, but it is not how I choose to live"). To conclude, disequilibrium serves to draw a person's attention to insufficient capacities as

1. More accurately, Lahey et. al. (1988) say that "If [a] speaker cannot demonstrate any knowledge of her being an active party to what she experiences, then we hypothesize that she is unable to do so, and that we have discovered her growing edge" (p. 37).

| authority-bound dualism | → | unqualified relativism | → | commitment in relativism | → | convictional commitment |

they attempt to remain in dialogue with a surround (the *immediate* down-turn in the helix above), and deconstruction serves to raze this way of knowing so it can be fully *rebuilt* into a synthesized, *new* frame of mind.

Figure 2. the process of deconstructing third-order to reconstruct as fourth-order as moving from

» DECONSTRUCTING THIRD-ORDER TO RECONSTRUCT AS FOURTH-ORDER

Let me now give a practical example of how this happens during a life passage common to American way of life: that delicate move from socialized thought to more institutional or systemic ways of knowing. A significant instance of explicitly questioning co-herence can come toward the end of socialized thought, or what Kegan (1994) might refer to as a shedding of the third-order way of knowing, to serve as an introduction to fourth-order thought. Third-order makes authority to be socially-based, it lies outside of the self. This frame is accomplished in that it does not entail a loss of self when needs or preferences are coordinated secondarily to perspective-taking. But its limits might be its subjection to external voices, and a lack of a chosen theory, or chosen way to make sense of experience. How might a third order way of know-ing undergo re-assembly? It usually begins with some confusion with regards to the previously rigid boundaries between "us" and "them." Sharon Parks (1986) calls this a reassessment of the very "Authority by which it was composed" (p. 48). This process might come about as the person begins to place trusted adults or peers, or entire disciplines or philosophical stances, within a relativized or pluralistic stance between one another.

Sharon Parks (1986), by condensing the models of William G. Perry (1970) and James Fowler (1981), refers to this process as moving from Authority-Bound Dualism through stages of relativism, and eventually to a Convictional Commitment (see Figure 2 below, from Parks, 1986)[2].

» AUTHORITY-BOUND DUALISM THROUGH STAGES OF RELATIVISM TO A CONVICTIONAL COMMITMENT

In this process, we see the person moving between authorities—one that is external and becomes reconstructed as internal. Through this reconstruction, the person-in-process will experience senti-ments similar to this young man interviewed by William Perry (1970): "Who will save me then from that 'wrong decision' I have been told not to make lest I 'regret-it-the-rest-of-my-life?' Will no one tell me if I am right? Can I never be sure? Am I alone?" (p. 33). Here we see clearly the tender limbo of transition. This young man is exploring the limits of the confines of his knowing, what was once strictly defined *for* him. But he has not yet recognized that he, in fact, will not be alone in his new frame of thought—al-though it may take some time for him to *re-cognize* the people he

now holds so dear (who he once conceived of as *holding him*). Soon this young man will find that it is he who makes decisions, it is he who defines what is worthy of regret, and it is he who coordinates his relationships. First he must place his dangling feet onto a new, unfamiliar platform that retrieves authorship or even a sense of stability out of the control of others and places it into his own hands. This allows a more sustained *revaluing* of these others or the concept of "others" as a whole.

This is the risk entailed in undergoing the rhythms of life. This young man's sense of isolation is a necessity, and could be seen as valuable. We rarely esteem isolation[3] but the developmental frame allows us to recognize that the isolation is a step toward some-thing else. The person is not *isolated*, but is *experiencing isolation* because of the reconstruction of his current meaning system. The above example serves as an introduction to an even more complex restructuring of meaning systems in the form of an entire field of study. But first, let me draw a working distinction between the self-authoring and self-transforming minds, to guide my analysis. According to Kegan (1994; p. 313), a self-authoring position sees the self-as-system as complete or whole, and identifies with the self-as-form interacting with other selves-as-forms. A self-transforming stance, on the other hand, regards self-as-system as incomplete, only a partial construction of all that the self is. The self-transforming order identifies with the process of form creation, bringing forms into being and subtending their relationship.

» CRITICAL DISCURSIVE PSYCHOLOGY AS POST-FOURTH ORDER

Before I analyze data from the field of critical discursive psychology, a history of its development is necessary. The birth of CDP comes directly out of cognitivist psychology, a study of speakers' accounts as reflecting their mental representations (Horton-Salway, 2001). The most significant theoretical contributors to its formation are Jacques Derrida (1991) and Michel Foucualt (1984), and involve a synthesis of linguistic philosophy, speech act theories, ethnometh-odology, conversation analysis, sociology of scientific knowledge, and post-structuralist study of cultural and literary texts. Instead of seeing language as a resource for evaluating inner events, discursive psychologists takes language as the topic itself, examining the ways social reality is constructed (Edley, 2001). Language, therefore, is not psychological *evidence*, but psychological *action*. Accounts are highly context-specific, and speakers are able to accomplish a variety of social actions through discourse. I will now address three key characteristics of critical discursive psychology as data, following each with a rationale for the ways in which it is evidence of a post-institutional way of knowing.

2. Some theorists (Parks, 1986; Gilligan, 1982; Fowler, 1981; Perry, 1970) depart from Kohlberg or Kegan by expanding the 3–4 shift into separate stages to account for the relativist period.

3. Westerners tend to value autonomy or independence—sometimes closely associated with isolation. I do not here value isolation, but rather advocate the recognition that a person is often in process, especially when voicing concerns similar to this young man's. It is not our role to adjust the person's discomfort, but it might be our role to ensure that the discomfort *continues to do its work* of re-establishing a deeper way of knowing.

» I. DISUNIFIED SELF

One of the most significant ways in which a critical discursive approach is deconstructive of the institutional way of knowing lies in its *denial of a unified self*. A modernist approach to psychology, and in some ways this includes the constructive-developmental approach, takes as implicit that a whole self, or a self striving for wholeness, consists of an inner life that is available for evaluation by the psychologist. On the other hand, Edley (2001) demonstrates that discursive psychologists and cultural anthropologists pose identities as "more fleeting, incoherent and fragmented than many of us would have believed" (p. 195). Indeed, the discursive analytic agenda is to reveal the Western "myth" of personal integrity and consistency of identity over time: contradictions in produced selves as objects will leave Westerners to feel defensive or embarrassed. Therefore, a discursive approach actively *seeks* variability, inconsistency, and variation-in-accounts—or, what have become called *ideological dilemmas* (Billig, 1988). Parker (1997) claims that this is not to "catch people out, but to lead us to the diverse and sometimes contradictory fragments of meaning that come together in any particular discourse" (p. 3). He roots this in Foucault's focus on contradictory discourses and the self as torn in different directions by these discourses. Deconstruction and its use of contradiction and variability, therefore, serves to dislodge the dominance of privileged forms of knowledge. Potter & Wetherell (1987) go on to assert that a multitude of selves is to be found in "the different kinds of linguistic practices articulated now, in the past, historically and cross-culturally" (p. 102). The material of self-as-theory-producer is considered culturally-prescribed. Ron Harré (1985) says that "to be a self is not to be a certain kind of being but to be in possession of a certain kind of theory" (p. 262).

Through a constructive-developmental lens, this stance is indeed a fascinating one. It is as if critical discursive psychologists integrate within the methodology a study of the very phenomenon addressed here: the deconstruction of favored or preferred, i.e. self-authoring ways of knowing! These ways of knowing are previously formed and become self-authored within a given society or subculture. For example, the client-centered therapy of Carl Rogers exhorts the person to drop façades and accept the whole self, and Gestalt therapy attempts to *recover* wholeness—for the purpose of discovering a true self, and finding the attainment to be in our own hands (Burr & Butt, 2000). These are clearly self-authoring guidelines. On the contrary, Parker (1997) asserts CDP to be going beyond this program by tracing distress to "networks of social relations and to patterns of language. Thus, a turn to discourse in therapy has helped therapists who want to link their work to wider issues of social justice" (p. 8). The hallmark of the self-authoring mind is the recognition that there is a cohesive self, a synthesis of different lives (social, professional, intimate, spiritual, etc.) which have been *chosen* rather than socially prescribed (Kegan, 1994). What makes critical discursive psychology of a more advanced frame is that it recognizes this, while going further to *sanction the disjointedness* of such a once-touted-as-united testimony or self-authorship. CDP is quite clearly a postmodern theory, and Kegan (1994) affirms that "what postmodernism is 'post' to is the fourth order of consciousness" (p. 317). The self-as-theory-producer referenced by Harré (1985) may be self-authored, but that self-authoring is realized to happen in the context of social material and reality, thereby making it only semi-self-authored or limitedly self-authored, and those limits are suddenly becoming explored by the person as discussed further below. Social material, therefore, are opportunities for a fifth order experience of "multiplicity" (Kegan, 1994). CDP clearly allows room for a paradoxical or post-ideological identity: a fifth-order accomplishment.

» II. PROCESS

The second element of critical discursive psychology under constructive-developmental scrutiny is its *emphasis on process*. Indeed, CDP seems to forsake psychological "position" for a capturing of linguistic *action*.

> *Discourse does not provide a transparent window into the mind of the individual or into the world outside, as many psychologists seem to believe. Rather, language organized through discourse always does things. When we seem to be merely describing a state of affairs, our commentary always has other effects; it plays its part in legitimizing or challenging, supporting or ironizing, endorsing or subverting what it describes. In both everyday language and in psychological description, our utterances are speech acts (p. 5, emphasis my own).*

In this account, Parker (1997) emphasizes the action orientation of speech. As discussed in preceding paragraphs, discursive psychology does not make claims to hold a person's inner experience as "data." What it does claim as "data" is *speech* located in discursive time and space, and considers the ways in which it accomplishes social acts. According to Edwards (2005), CDP identifies how people are "shown to formulate or work up the nature of events, actions, and their own accountability through ways of talking. These ways of talking are both constructive and action oriented" (p. 260). Thus, studies are conducted on event reporting, script formulation, the use of emotional syntax in personal narratives, and other discursive activities or, more accurately, *accomplishments*. Michael Halliday (1985) calls this "transitivity," holding basic grammatical categories through which language of agency and action construct a social world full of events and social actors that perform such feats as owning, obscuring, downplaying, or omitting agency.

Robert Kegan (2010) finds a post-ideological formation to be highly involved with movement, process, or modification rather than categories, stasis, self-completeness, or imperatives. Here we find another break from a former (modernist) emphasis on psychological location. For example, Jean Piaget's genetic epistemology valued such (fixed) things as structured whole, end, means, quality, quantity, object, space, causality, and so forth (Morra, S., Gobbo, C., Marini, Z, & Sheese, R., 2008). But because CDP is interested in social processes, linguistic and ideological movement gives rise to contradictions in discourse. As mentioned above (and below), CDP centers its focus on contradiction. Motion and oscillation between paradoxical stances are crucial. Michael Basseches (1984) confirms that for the self-transforming (dialectical) mind, contradiction is not only viewed as positive, but necessary. "Every system," he says, "is limited in its ability to maintain stability" (p. 123). As such, forms and entities are found as limiting, but motion, process, and change are the focus of study for discursive psychology and indeed prioritized in a fifth order framework (Kegan, 1982).

» III. SUBJECTIVITY

The last element of critical discursive psychology to be analyzed here is its self-criticism as a theory-making theory. CDP readily admits its own subjectivity while operating within a discipline it criticizes as socially-constructed (psychology). Indeed, subjectivity arises from this assertion of socially-constructed knowledge, as Potter & Wetherell (1987) elaborate…

> *The assumptions on which modern Western psychology is based, the concept of the self as the center of experience, for instance, which we discussed above, may be peculiar to this period of history and this type of society. What psychology discovers, therefore, may not be the timeless universal features of personhood. Psychology may simply elaborate instead upon the conventional ways people are described in this particular society (pp. 102–3).*

As it identifies its limitations, it is also actively seeking disciplinary partners particularly as it was born out of cross-disciplines, as noted above in the historical review of CDP. Discourse and stakes of power occur in the day-to-day. It is here that Parker (1997) identifies promptings for a need for discursive analysis, but only as it is affiliated with "life experience and political identities *outside* the discipline. Only then does it make sense to deconstruct what the discipline does to us and to its other subjects" (p. 13, emphasis in original). For Parker, this involves addressing how psychological facts are socially constructed, how subjectivity is discursively reproduced within present social arrangements, and how the underlying historical conditions emerged that gave rise to the psycomplex. Understanding how the field of psychology reproduces realist notions like individuality and human nature opens the possibility to "transform it, and to socially construct it as something different, something better" (p. 11). A theory closely aligned to CDP is Grounded Theory, which attempts to construct theory out of everyday discourse, "grounding" it closely with the data rather than imposing a pre-formed theory on the data. This self-criticism and grounded reconstruction begins with a strategy in the field known as "subjectivity." Whereas scientific (modernist) research methods heralded objectivity, a discursive approach not only claims subjectivity but *values* it. A constructed theory includes such subjectivity, and Kathy Charmaz maintains that we are all—even as researchers—part of the data that is collected (2006). As a whole, discursive psychology is both self-aware, and self-proclaimed as undergoing its own process of formation (Edwards, 2005). It appears that the theory itself, with its prizing of process, reflexivity, and identity-as-fragmented, makes room for such lack of wholeness in its own architecture.

Critical discursive psychology could be seen as fourth-order in that it is theory-producing. It makes claims that seem absolutist or all-encompassing. Yet the theory takes as central to its formulation the capacity for "self-"criticism, and identifying the limits of its own theory. Belenky, et. al. (1986) see fourth order, "procedural knowing," as being surpassed by an epistemology of "constructed knowing." Here, we emerge from our self-authorship more aware of and *responsible for* the constructed nature of our realities. Kegan (1994) affirms this: "A theory that is also a theory about theory-making is about the *way* we take a stand will by necessity make judgments about and deprioritize those procedures, theories, and stands that are not self-conscious about their own tendency toward absolutism" (p. 330; emphasis in original). This appears consistent in a discursive psychological approach, with its emphasis on dilemmatic formulations of ideology and contrary elements of thinking, rather than treating ideological systems as integrated ways of knowing, "as schemata *par excellence*" (Billig, et. al., 1988; p. 27). CDP mirrors the dialectical self in that its formulations "affirm a bipolarity, opposition, or contradiction as the fundament of dynamic systems" (van den Daele, 1975; p74). Indeed, CDP is consistent with Arlin's (1975) proposal of a fifth-order *problem finding*. She claims that while "formal structures are stable, having achieved equilibrium and remaining available throughout life," they become surpassed by creative thought characterized by problem-discovery processes (p. 602). Although she fails to show here how this problem-finding becomes directed toward one's own stable or achieved "formal structures," she duly emphasizes that which CDP is about: a respect for inconsistency and not for the purpose of providing wholeness.

On the topic of subjectivity, discursive analysts are able to comprehend that ideologies form as a means of making sense of a complex world. We either unknowingly blend into pre-formed narratives (3rd order), or form our own narratives whether they be pre-existing or not (4th order). But a 5th order stance recognizes the *limits of preferred ideologies*. Although necessary, ideologies are also restricting. Critical discursive psychology edges closer to a firmly-rooted fifth level plateau in that it does not problematize the holding of a preferred ideology as a deconstructive post-fourth order, or anti-modernist[4] might (Burbules & Rice, 1991). It attempts to bring out as much as possible, and cherish, subjectivity through a process of memoing (Charmaz, 2006). This process allows an opportunity to evaluate what the research knows, knows she doesn't know, and doesn't know she doesn't know (Kegan, 2010). CDP is fifth order, indeed beyond an anti-modernist emphasis on deconstruction, in that it *re*-values preferred ideologies but seeks to identify the certain boundaries they carry, which is important because it instructs us and the researchers on the limits of preferred theories. Overall, subjectivity and theory re-producing serves as a reformulation of the previous institutional way of knowing, as the fifth order reorganizes and acquires these concepts for use in new ways and to serve new functions (Turiel, 1974).

To conclude this section on CDP emerging as fifth-order, I will review what Kegan (1982) calls the "cracking" of the institutional self. The construction of "the self as system, form, or institution of which 'I' am the administrator who must keep the organization intact, *a way of seeing now seen through*" (p. 237). Deconstructive tendencies of stage transition show how the person has not yet integrated the former-form into a new form. Critical discursive psychology is a re-conceptualization or more complex recasting of cognitivist psychology rather than a splinter-theory or reactive

---

4. For Burbules and Rice (1991), antimodernism (as opposed to a *fully* post- or reconstructive modernism) upholds the deconstructive strain addressed in the first half of this paper. It is not "concerned with recapturing and reformulating modern values, such as reason or equality, but with deconstructing them. Not surprisingly, this tradition in particular has been more convincing in pointing out the limitations and contradictions of modernism than in reformulating positive alternatives" (p. 398).

theory, thereby re-legitimizing its preceding form of study. CDP adopts the fifth order understanding that ideologies only provide a partial view of reality (Billig, et. al., 1988)—a reality that is complex, moving, and paradoxical. They must be deconstructed so as not to be left disassembled, de-institutionalized, and removed of power, as an antimodernist might offer; but deconstruction allows a reformulation (a *re-form*-ing) of abstract systems to appreciate their limitations, constructed-ness, and divisions which give them their identity, fully validating their potency *and* futility.

» CONCLUSION

Because all transitions involve leaving behind a consolidated self to make space for a new self, the disruption in self-unified-theory can be disharmonious. It often entails facing an abandonment of or denial of reliance upon a form, group, standard, convention, indeed a very system of meaning-making (Kegan, 1982). Checking myself against my own disclaimer above about over-simplifying complex transitions, we must conceive of this transition as a jarring demolition of what might otherwise feel like to be a refuge, in the self-authoring stage. Kegan (1982) talks of this transition as getting the sense of "leaving the moral world entirely" and disorientation regarding right and wrong or the institutional conception of a standard (p. 232). And how might this feel for the context—the holders of the person-in-process? It might feel like betrayal, faithlessness, disloyalty, deception, or even signs of mental illness. But we must recall the developmental hope that Kegan (1994) offers: "*differentiation always precedes integration*" (p. 326). Before an integration of previous forms can take place, distance must be maintained. Parents will take personally a two-year-old's first signs of protest in their increasing use of the word "No." But Kegan (1994) reminds us that this is "literally its first objection, a declaration that it is making into object the people and things with which it had formerly identified itself" (p. 326). So it may not so much be the parents this child is objecting, but the child's former organization of the self.

The institution that experiences a barely post-fourth order objection often reacts vehemently at the deconstructions of its heralded truth, only perpetuating ideological wars. But the fourth order institution, and the emerging fourth order questioner-of-coherence, will forget that such a stance is often untenable and will likely give way to a re-absorption and re-cognition of institutional values and limits. Anti-modernists, therefore, can look to critical discursive psychology as an advisor in the ways of not just deconstruction, but healthy reconstruction. Its ability to re-integrate institutional characteristics, as well as its former way of knowing—cognitive psychology—give it a more stable stance in the fifth order of making meaning.

William G. Perry is reported to having had a favorite axiom: "The person is always larger than the theory" (Parks, 1986; p. 41). A person-in-process is especially larger than any theory, and will experience discomfort that can hardly be appeased by insight from a theory. Therefore, in my concluding thoughts, I turn to an excerpt of an interview between bell hooks and Pema Chödrön. In discussing the necessity of death, Chödrön asserts that

"For me the spiritual path has always been learning how to die. That involves not just death at the end of this particular life, but all the falling apart that happens continually. The fear of death—which is also the fear of groundlessness, of insecurity, of not having it all together—seems to be the most fundamental thing that we have to work with. Because these endings happen all the time! Things are always ending and arising and ending. But we are strangely conditioned to feel that we're supposed to experience just the birth part and not the death part (bell hooks, 2010, p. 3)."

Seasons come, and seasons go. The lush shrubs of summer make way for death and rotted leaves, and eventually naked trees. And darkness; the cold, bitter isolation of long winter nights. But soon nights get shorter and new sprouts re-emerge. All people will die; all humans-being are in a process of death and decay. Cells join, but so too do they divide. Beyond just physical deterioration, all humans feel loss and instability to various degrees and for various periods of time. Chödrön's advice, or promise—indeed, a constructive-developmental[5] "truce," to use Kegan's (1982) early term for harmonious equilibrium—is to avoid the struggle against struggle. As this paper has affirmed, deconstruction or Chödrön's "struggle against struggle leads to new integration and composition of a whole new way of knowing, or a wider view of reality. ∎

5. Chödrön seems especially constructive-developmental earlier on in the interview when she asserted that "Getting stuck in any kind of self-and-other tension seems to cause pain."

REFERENCES

Albertson, S. (2010). Growth and construction: The subject-object dialogue. The 25th Annual Adult Development Symposium, Society for Research in Adult Development. Philadelphia, PA. 10 March 2010.

Basseches, M. (1984). *Dialectical thinking and adult development*. Norwood, NJ: Ablex.

Basseches, M. (1989). Toward a constructive-developmental understanding of the dialectics of individuality and irrationality. In D.A. Kramer & M.J. Bopp (Eds.), *Transformation in clinical and developmental psychology* (pp. 188-209). New York, NY: Springer-Verlag.

Belenky, M.F., Clinchy, B.M., Goldberger, N.R., & Tarule, J.M. (1986). *Women's ways of knowing*. New York: Basic Books.

Billig, M., Condor, S., Edwards, D., Gane, M., Middleton, D. J. & Radley, A. R. (1988). *Ideological dilemmas: A social psychology of everyday thinking*. Thousand Oaks, CA: Sage.

Burbules, N.C. & Rice, S. (1991). Dialogue across differences. *Harvard Educational Review*, *61(4)*, 393-416.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: Sage.

Derrida, J. (1991). *A Derrida reader*. P. Kampuf (Ed.). New York, NY: Columbia University Press.

Duckworth, E. (2006). Piaget rediscovered. *Journal of Research in Science Teaching, 2(3),* 172-175.

Eagleton, T. (1987). *Frère Jacques: The politics of deconstruction*. Berlin, DE: Mouton de Gruyter.

Edley, N. (2001) Analyzing masculinity: Interpretive repertoires, ideological dilemmas and subject positions. In M. Wetherell, S. Taylor, & S. Yates (Eds.), *Discourse as Data* (pp.189-228). Thousand Oaks, CA: Sage.

Edwards, D. (2005). Discursive psychology. In K.L. Fitch & R.E. Sanders (Eds.), *Handbook of Language and Social Interaction* (pp. 257-273). Mahwah, NJ: Lawrence Erlbaum Associates.

Foucualt, M. (1984). *The Foucault reader*. P. Rabinow (Ed.). New York, NY: Pantheon.

Fowler, J.W. (1981). *Stages of faith: The psychology of human development and the quest for meaning*. San Francisco, CA: Harper & Row.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.

Ginsburg, H.P., & Opper, S. (1988). *Piaget's theory of intellectual development* (3rd ed.). Englewood Cliffs, N.J.: Prentice-Hall.

Halliday, M.A.K. (1985). *An introduction to functional grammar*. London, UK: Edward Arnold.

Harré, R. (1985). The language game of self-ascription: A note. In K.J. Gergen & K.E. Davis (Eds.), *The Social Construction of the Person* (pp. 259-263). New York, NY: Springer-Verlag.

Hooks, B. (2010). News you can use: Pema Chödrön & bell hooks talk over life and all its problems. Originally printed in Shambhala Sun magazine. Retrieved from Shambhala Publications: <http://www.shambhala.com/html/learn/features/pema/interview/index.cfm> 09 May 2010.

Horton-Salway, M. (2001). The construction of M.E.: The discursive action model. in M. Wetherell, S. Taylor, & S. Yates (Eds.), *Discourse as Data* (pp. 147-188). Thousand Oaks, CA: Sage.

Kegan, R. (1982). *The evolving self: Problem and process in human development*. Cambridge, MA: Harvard University.

Kegan, R. (1994). *In over our heads: The mental demands of modern life*. Cambridge, MA: Harvard University.

Kegan, R. (2010). Life beyond 'self-authoring:' Exploring the fifth order of consciousness. Harvard Graduate School of Education, Adult Development. Cambridge, MA. April 12th.

Lahey, L.L., Souvaine, E.A., Kegan, R., Goodman, R., & Felix, S. (1988). *A guide to the subject-object interview: Its administration and interpretation*. Laboratory of Human Development. Cambridge, MA: Harvard University.

Lewis, M.D. (1995). Cognition-emotion feedback and the self-organization of developmental paths. *Human Development, 38(2)*, 71–102.

Maton, A., Hopkins, J, Johnson, S.J., LaHart, D., Warner, D.Q., & Wright, J.D. (1997). *Cells: Building blocks of life*. Upper Saddle River, NJ: Prentice Hall.

Montanegro, J. (1985). *Genetic epistemology: Yesterday and today (Pro Helvetia Swiss lectureship).* New York, NY: The Graduate School and University Center.

Morra, S., Gobbo, C., Marini, Z, & Sheese, R. (2008). *Cognitive development: Neo-Piagetian perspectives*. New York, NY: Psychology Press.

Parker, I. (1997). Discursive psychology. In Fox, D. & Prilleltensky, I (Eds.), *Critical Psychology: An Introduction* (pp. 284-298). Thousand Oaks, CA: Sage.

Parks, S.D. (1986). *The critical years: The young adult search for a faith to live by*. San Francisco, CA: Harper & Row.

Perry, W.G. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. New York, NY: Holt, Rinehart, and Winston.

Piaget, J. (1974). *Adaptation vitale et psychologie de l'intelligence: Selection organique et phenocopie*. Paris, FR: Hermann.

Piaget, J. (1985). *The equilibration of cognitive structures*. Chicago, IL: University of Chicago Press.

Potter, J. & Wetherell, M. (1987). *Discourse and social psychology: Beyond attitudes and behavior*. Thousand Oaks, CA: Sage.

Selman, R. L. (1980). *The growth of interpersonal understanding: Developmental and clinical analyses*. New York, NY: Academic Press.

Turiel, E. (1974). Conflict and transition in adolescent moral development. *Child Development, 45*, 14-29.

Van den Daele, L.D. (1975). Ego development and preferential judgment in life span perspective. In N. Datan & L.H. Ginsberg (Eds.), *Life-Span Developmental Psychology* (pp. 1-67). New York, NY: Academic Press.

Zizioulas, J. (1993). *Being as communion*. Crestwood, NY: St. Vladimir's Seminary Press.

BEHAVIORAL DEVELOPMENT BULLETIN

Leadership and adult development:
Towards a unified neuro-psycho-economic approach

Marc G. Lucas[1] and Svenja Caspers[2]

1 University of Hagen
2 Institute of Neuroscience and Medicine (INM-1)

ABSTRACT

The paper first summarizes basic findings of the ongoing interdisciplinary research project on differences in neural processing of individualistic vs. collectivistic oriented test persons (managers vs. non-managers). Test persons had to perform abstract moral decisions within a functional magnetic resonance imaging (fMRI) setting. The obtained neurobiological and behavioral data were compared between the above mentioned extreme groups.

In this paper the integration of dominant psychological trait theories (BIG-5) and theories of adult development (AD) will be first established on a theoretical level via a synopsis of Furnhams (1996) findings on the integration of different psychological trait theories and neuropsychoeconomic dual process theories primarily related to the importance of intuition in decision making (Kahneman, 2003). A characteristic pattern of a combination of traits will be presented as a possible marker for a high System 1 activation. This pattern will be tested as a signifier for a concording higher development in AD as well. This theoretical approach will be validated by empirical data from the project in which the researchers combined the extreme group analysis with the application of psychological tests as e.g. NEO-FFI and WUSCT.

KEYWORDS: neuropsychoeconomy, individualism vs. collectivism, fMRI, BIG-5 personality traits, dual process theory, WUSCT

NEUROPSYCHOECONOMICS DESCRIBE the human behaviour and experience in economic context with methodological support of neurosciences (Camerer, et al., 2005; Schilke & Reimann, 2007). As transdisciplinary science, it combines theoretical input in particular from psychology, sociology and economic sciences. According to Küppers (2000), it also brings together a broad range of approaches from social sciences and natural sciences, while structural sciences (mathematics, IT and logic) constitute the link. Hence, it approaches the biopsychosocial paradigm of metascience and beyond this it also facilitates empirical research for a broader understanding of empiricism. The useful addition and mutual enhancement of traditional qualitative and quantitative survey and behavioural observation procedures with methods of brain research in the sense of a triangulation lies in the forefront of the neuropsychoeconomical research design. In the field of neuroscientific methods, functional magnetic resonance imaging (fMRI) has distinguished itself in particular. This imaging method allows for measurement of oxygenated blood in the brain using strong magnetic fields. This activation measurement takes place

continuously and enhances while performing activities. Particular insights are hence obtained from precise measurement of parts of the brain which are active at the time of certain activities, versus those which are not involved or inactivated at the time of these activities. Starting with Fischer (1971), the knowledge obtained in the field of neurosciences has also been applied in the research of changed (or higher) states of consciousness. Here, the research with practitioners of meditation has distinguished itself in particular. It shall only be mentioned in passing that here, research is carried out only on the non-ecstatic and at the same time inactivated dimension from the two dimensions of hyperarousal and hypoarousal of higher states of consciousness which have been described by Fischer.

Neuropsychoeconomics primarily focuses on the contribution of new emotion theories to the decision-making behaviour especially of executives. A decision-oriented and neuropsychoeconomically based definition of leadership is provided by Lucas (2012, p. 9; 2013): "Leadership can thus from an integrated multiperspective view be broadly defined as a specific personality, socially and neurologically caused targeted action initiation under pressure to decide, which is at the same time retroactive for these spheres." In a neuro-psycho-economic research par-

Author's note: correspondence regarding this article should be addressed to Marc G. Lucas, Ph.D., Institute of Leadership and Organization, Department of Economics University of Hagen, Hagen Germany, marc.lucas@fernuni-hagen.de

adigm, this approach and definition is being followed through differentiated measurement of the activity in certain regions of the brain and crossparadigmatic convergences are being looked for. Preoccupation with the significance of different activity patterns begins already with the purely psychological circumplex theories of emotion and personality which have a wide scope. Theories are presented here which aim at presenting the different dominant traits of emotions as well as the Big 5 in the field of personality research in a horizontal circle (Gurtman, 1997, p. 83–84; Plutchik, 1997, p. 21–22). Starting with Watson and Tellegen (1985, 1999), circumplex models are increasingly being presented which are based on two independent activation systems in emotion regulation. In particular Schallberger & Pfister (2001, p. 179f) as well as Schallberger (2006, p. 98f) have transferred this representation of positive and negative activation to the context of leadership and to management decisions as well. Csikszentmihalyi (cited according to Aellig, 2004, p. 44) uses this approach for explaining his flow concept as a highly creative and modified state of consciousness of positive activation in absence of negative activation.

Genuinely neuropsychoeconomical theories deal with these insights and examine these within the framework of theories regarding bounded rationality of decisions. Dismissing the key economical premises, they emphasise that an individual does not decide on a purely rational basis and "that the homo economicus lacks an emotional basis" (Peters & Ghadiri, 2011, p.11). Among the dual process theories there, the approach of Kahneman (2003; 2011) has distinguished itself in particular. He believes that humans make decisions on the basis of two basically different but potentially concurrent processing systems in the brain. In his System 1 (in subcortical-emotional processes, the region of the dorsal striatum is involved), the arriving information is rapidly and intuitively processed. Simple heuristics and the pattern recognition processes lie at the forefront of the processing. On the other hand, System 11 (cortical-cognitive) assesses information extensively and in a balanced way in order to finally arrive at an adequate decision. The arriving stimuli pass through different cortical perception and cognitive weighting filters which correspond remarkably to the ego-stabilising filter processes mentioned in Loevingers' (1976, 1993) theory of ego-development. This considerably slower system can therefore be termed as a rational system as well. Quite astonishingly, the strategies of System 1 proved to be equally successful at least if not better than System 11 which was initially considered to be more accurate.

These hypotheses and initial findings from the field of neuropsychoeconomics received support in the meanwhile from other empirical studies. For example Greene et al. (2001, 2004) and Koenigs et al. (2007) also proved in fMRI studies that in processing of moral dilemmas and complex tasks, cortical cognitive-rational as well as subcortical emotional process components play a role. It must be highlighted however that rapid processing in System 1 is not exclusively emotionally controlled. Emotions only have a certain influence on it. Hence, emotions are not the only factor that affects intuitive information processing in System 1 according to Kahneman (Fischer & Wiswede, 2009) but are linked to personality-related factors as well.

The fMRI based research on the dual process theories is of high significance for adult development psychology with respect to its possible contribution. The dual process theories focus on the high significance of non-rational or post-rational intuitive decision-making processes in successful addressing of complex tasks and the decision regarding moral dilemmas. Refer to the discussion on the empirical status of the post-formal development level (Commons et al. 1995; Marchand, 2001) for elaboration. Similarly, the theories and methods available in Adult Development research for measurement of later formal and post-formal development levels could be valuable for crossparadigmatic neuropsychoeconomic research. However, the complete absence of an approach integrating these research fields must be noted. Only Haidt (2001) deals with the postulate of a purely cognitively controlled assessment of morally relevant situations as formulated by Kohlberg (1969). He also reaches the conclusion that in presence of increasing neuroscientific knowledge, this hypothesis is no longer sustainable. In 2003, Kallio & Pirttilä-Backman (p. 137) stated that: "A holistic approach is also needed to achieve an integration of distinct, often separate, research subjects (i.e., cognition, personality, emotions)." Even if we are aware that "research across domains has been inhibited by theoretical and practical problems" as formulated by Commons et al. (1989, p. 33), the initial explorative steps in directing integration of the above-mentioned research fields to a neo-integral theory should be taken with the existing and ongoing work of Caspers et al. (2011, 2012).

» THEORETICAL FOUNDATION

The early and trend-setting insight of Graves (1970) about the necessity of integration and a neurobiological basis of adult development theories within the framework of a bio-psycho-social paradigm was the starting point for the model-theoretical considerations and the research design. In his hierarchical and open-ended development theory, intra-psychic development of the individual (labels A–H) and social development of collective life conditions (labels N–U) are mutually dependent. He has differentiated eight levels in his "theory of human existence levels" at current time which are further divided into six subsistence and two being levels of second order. Each subsistence/being level is seen in humans through a corresponding system of values that characterises their entire perception and behaviour.

Switching between the levels leads to higher complexity and is a result of the fact that a person realizes a dissonance between the perception of "himself" and the environment. Ideally, a distinction is made thereby between development levels in which a human being adjusts his environment to his needs and those levels in which he adapts his needs to his environment. Development takes place in switching between these rather collectivist and rather individualistic systems of values. The characteristics of these levels can be described with regard to various individual and social structures. In table 1 some fundamental values have been described which would predominate in leaders with concentration on one of the eight postulated levels. In the meantime numerous empirical and theoretical studies support the assumption, that life- and work-conditions of leaders are characterized by higher complexity (Beer, 1995; Jacques, 1998) especially in the field of decision making. In consistency with the expectations of the theory of levels

Table 1. levels of existence (LOE) with corresponding leadership values (in dependence on Graves, 1974, Beck & Cowan, 2007 and Hamilton, 2012)

level of existence	basic leadership values
A−N (beige)	*reactive values:* survival; staying alive through innate sensory equipment.
B−O (purple)	*traditionalistic values:* belonging to blood relationships, safety, assurance
C−P (red)	*exploitive values:* enforcing power over self, others, and nature through exploitive dependence, egocentrism
D−Q (blue)	*sacrificial values:* commanding absolute belief in one right way and obedience to authority, salvation, order, security, absolutism
E−R (orange)	*materialistic values:* inventing and possibility thinking, focused on making things better for self. rationalism, multiplism, independence
F−S (green)	*relativistic values:* sharing equality and seeking the well-being of people; building consensus as highest priority, community, affiliation
G−T (yellow)	*existential values:* adapting flexibly to change through connected, systemic views, self-worth
H−U (turquoise)	*experientialistic values:* attending to whole-earth dynamics and macro-level actions, communion

of human existence leaders on an individual level generally show centres of gravity in later adult developmental levels also (Smith, 1980; Merron, 1985; Quinn & Torbert, 1987; Corbett, 1995; Rooke & Torbert, 1998; Cooper, 2005; Brown, 2012).

Even if the "theory of human existence levels" remained largely unknown in academic circles due to the early death of Graves and little empirical research has been published till date, it has won a lot of interest just lately. For example, Strack (2011, p.22) has published an empirical study in which evidence is found, that Graves' theory and Schwartz's typological theory of the universalistic value circle (Schwarts & Bardi, 2001, Schwartz, 2006) which is dominant in today's value research are based on the same value structure.

Graves et al. tried to describe a neurobiological correlate of his theory already in 1965 using a tachistoscope. In this study, specific value terminology was presented to the test persons for a few milliseconds at a time. Terms that corresponded to the dominant effective value system of the test person were recognised more quickly than other terms.

This approach was basis for our current studies, which transferred the original idea into an up to date research design backed by FMRI and complemented by numerous psychological procedures (CFT intelligence test, NEO-FFI personality test, Values Test and WUSCT Ego-Development).

» PREVIOUS STUDIES
Study 1:

In the initial study (Caspers et al., 2011), 38 healthy German-speaking test persons (21 male, average age = 39.7 years, 17 female, average age = 33.7 years) were selected from a larger test group. They were asked to choose between value-laden abstract term pairs within the framework of a forced-choice FMRI research design. The terms were partly taken from the original study of Graves and partly from the Values Test, which is a questionnaire-based method based on the theory of Graves. In addition to this, terms were developed with respect to key issues of the ego-development theory which were assigned to the levels set by Graves. This addition was made to allow for orientation to a better understood adult development theory with validated procedures. But it may have also been because of this composition that on basis of FMRI data, only classification of different processing patterns between individualistic and collectivistic value concepts could be established. Such a distinction was also recognisable between subsistence and being levels related value terminology. But the distinction did not prove to be significant due to the low number of test persons which could be explicitly assigned to the last levels.

Study 2:

In a second study (Caspers et al., 2012), 44 healthy German-speaking executives (22 male, 22 female, average age = 44.3 years) from the fields of science and economics and 35 non-managerial staff (average age = 40.4 years, 23 male, 12 female) were examined through FMRI in the same study design. From the 44 healthy German-speaking executives, 35 could be considered for the assessment. The non-managerial staffs were chosen from a large sample of 82 test persons who took part in other functional studies. Here, differences were seen in recruiting of brain areas in terms of dominance of the rational or intuitive processing system as assumed on basis of dual process theories. The executives in the study used a subcortical region within the basal ganglia system (head of caudate nucleus) which is thought to gate the way for automated processing of stimuli. On the other hand, the non-managerial employees used a large network of cortical brain regions, which speaks for the use of the rational processing system II. In this network, areas for visual processing as well as for decision-making processing are involved according to the visually presented stimuli.

» CURRENT STUDY—DESCRIPTION OF SECONDARY ANALYSIS AND HYPOTHESES

This knowledge is now considered as an opportunity to attempt the integration of the following within the framework of a secondary analysis of the data available from both studies:

1. the personality theory of the Big-5 measured through NEO-FFI with focus on a specific pattern for intuition,

2. the behavioural data related to the choices made during execution of the FMRI paradigm and

3. the adult development models (human existence level according to Graves which is measured with Values Test as well as ego-development according to Loevinger which is measured with the Washington University Sentence Completion Test).

The first indication of this integrative approach is provided by some descriptive studies. These descriptive studies adopt the difference between managerial staff and non-managerial staff in *a)* attribution of leadership qualities (Judge et al., 2002), *b)* in the management efficiency and the likeliness that they assume executive positions (Ilies, Gerhardt & Le, 2004, p. 215) and *c)* fundamentals of transformational leadership (Weibler, 2012, p. 112) in the Big-5 personality dimensions. In particular, for the

dimensions "Neuroticism"(negative attribute) and "Extraversion", and conditionally also for "openness "and "Conscientiousness", significant correlations have been reported. For "Agreeableness", no such effect has been found.

If one wants to bring together these findings in form of an explanatory approach, the reported personality-oriented markers of executives can be brought into relation with the result of previous study 2. According to study 2, executives activate the intuitive system 1 in a large number of cases in an abstract decision-making situation. Furnham (Furnham & Stringfield, 1992; Furnham, 1996) managed to work out a more extensive description of "intuition" as a specific pattern of the Big-5 in a comparison of the Big-5 personality dimensions. The dimensions were "Extraversion-Introversion" and "Sensing-Intuition" of Myers Briggs Type Indicators (MBTI). Furnham also associated these dimensions with leadership qualities. Thereby he recognised "Neuroticism" as a general characteristic of negative emotionality or of increased potential for negative activation which at high levels can lead to wrong decisions in the sense of aversion and stress reactions. A pattern of low neuroticism, high extraversion and high openness could be described as "intuitive" and a pattern of low neuroticism. But low extraversion and high conscientiousness at the same time could be described as "largely perceptual" as a specific personality trait. A mixture of both patterns can only be seen in case of middle levels of extraversion.

The association of intuition with leadership qualities or with the executives, mentioned in our former study 2 or in Furnham's study, suggests that the activation pattern described as intuitive system 1 in the dual process theory of Kahneman has its counterpart in the pattern consisting of personality traits. This was described by Furnham as "intuition" and the rational system 11 according to Kahneman. The rational system 11 is associated in our studies with significant activation of cortical areas of perception processing. It has its counterpart in the personality pattern of "largely perceptual" in case of Furnham. If one adds the significance of pattern recognition and intuition for higher levels of adult development, it may be assumed that the described personality traits of intuition and system 1 activation should also correlate with more complex levels of adult development.

This leads to the central hypotheses of this study:

Hypothesis 1: Executives show a higher level of development in adult development theories. This is due to selective processes and tasks which are characterised by higher complexity. The executives also make corresponding selection decisions in the forced-choice FMRI design, which show a preference for higher existence-level values.

Hypothesis 2: Executives exhibit a personality pattern which is closer to the "intuitive" type. Non-managerial staff members have a personality pattern which is closer to the "largely perceptual" type.

Hypothesis 3: The personality pattern of the "intuitive"-type correlates with a higher level of adult development and corresponding choice decisions made during execution of the FMRI paradigm.

Hypothesis 4: The "largely perceptual" personality pattern correlates with the intermediate forms of adult development and the corresponding preference of the choices made during execution of the FMRI paradigm.

Hypothesis 5: The personality trait of "neuroticism", as a trait of aversive affectivity and overload through complexity and decision-making under pressure, correlates with a low grade of adult development and the preference for choices made during execution of the FMRI paradigm.

The initially described development theory of human existence levels does not primarily and exclusively relate to individual traits. Similarly, no adequately validated methods are available for this theory. Only a study of Karsten (2006) deals with the Values Test and draws the conclusion that this method has inadequate internal consistency. Karsten also concluded that it was difficult for the test persons to understand the formulations of the questionnaire or to correctly perform the repeatedly requested allocation of 15 points to seven dimensions. Therefore, a better established, projective and hence bias-free and at the same time better validated adult development theoretical procedure (Loevinger, 1970, 1976) was implemented. This development theoretical procedure was implemented in the form of the Washington University Sentence Completion Test (WUSCT) which is oriented to the theory of ego-development. This procedure has been implemented in several leadership-oriented studies (Lucas, 2012, p. 13, Lucas, 2013) and according to Robinson (2013, p. 157) can be attributed to the adult personality development approaches. For scoring, an analysis could only be carried out by a trained rater by way of derogation from the procedural guidelines.

» RESULTS

Sample characteristics: The combined sample that is presented here comprises the data sets of Studies 1 and 2. It was enhanced by the extended dataset of the 82 preliminary subjects from the Forschungszentrum Jülich in the Helmholtz Association. After excluding one incomplete dataset, this produced a combined sample of 125 people. In accordance with the guidelines we had set for Studies 1 and 2, these were equally divided between women ($n = 61$) and men ($n = 64$). The distribution was also almost equal with respect to leadership function—defined as having leadership responsibility for at least 5 employees—(leadership function = 52, no leadership function = 73). The slight preponderance of subjects without leadership responsibility was due to the fact that this selection criterion was not applied in the preliminary study. However, the leadership and employee functions are equally distributed between the sexes.

The presence of WUSCT protocols, which can be attributed to the postformal stages, is of particular interest and is therefore displayed separately in these results. Unfortunately, as there was no pre-selection process for participating in the study based on such ratings, even when including later protocols of the individualistic stage (4/5) it was only possible to identify 4.8% of all participants in the combined sample with postformal ego identity. In absolute figures this corresponds to just six cases, which unfortunately makes it impossible to provide statistical evidence.

The data from the FMRI study could be taken into account in so far as it was possible to classify the individual subjects as individualists or collectivists using the choice pattern made during execution of the FMRI paradigm. This produced an almost equal distribution, with a total of 69 individualists and 56 collectivists.

Table 2. descriptive statistics of current study (gender, leadership, postformal stage and neurosplitgroup)

	total sample	female	male
number	125	61	64
leadership *N*	52	25	27
no leadership *N*	73	36	37
postformal stage *N* (beginning WUSCT late individualist)	6	2	4
neurosample individualist	69	31	37
neurosample collectivist	56	30	26

There were only minimal differences between men and women, with a slightly stronger individualistic tendency in men and a slightly stronger collectivist tendency in women.

At 42, the average age of the subjects was quite high. This was because participation in the study was restricted to adults and because of the large number of leaders in the study. And as the study was focused solely on people in employment, the oldest participant in the study was 61 years of age. The participants had an average IQ of 124 as measured by the Culture Fair Intelligence Test (CFIT 20). This high level was due to the fact that some of the subjects were acquired via Mensa, but it also reflects the high proportion of leaders in the study. However, a statistically significant correlation to the level of adult development could not be identified for the ego identity or for the levels of existence. With respect to the tendency towards particular personality traits in the intuitive type and the perceptive information processing type, there was no evidence of a correlation to IQ levels.

After evaluating an abbreviated version of the WUSCT with 18 sentence stems, a small statistical spread of overall ratings was established. The range included only five levels with the sample size being too small at the postformal stages for a statistical evaluation. The focus of the protocol was on the Self-Aware (*n* = 68) and Conscientious (*n* = 29) stages. A problem for the evaluation was posed by the often very short answers given by the subjects, something that may have been due to the fact that participants not only had to fill in detailed questionnaires for this study but also other questionnaires for a parallel study.

The levels of existence were measured using an abbreviated version of the Values Test, consisting solely of the preferred choices for the individual stages of the theory. The second part of the test, which included the same questions in the form of degrees of rejection, was not used for practical reasons and because of the aforementioned comprehension problems experienced by the participants. A wider statistical spread was noted across the eight levels of existence. However, the range of centres of gravity that were allocated to the individual subjects was spread across all eight levels of existence. These major differences between the two procedures for measuring adult development can to some extent be explained by the relative independence of the underlying construct. However, the different ways of evaluating the procedures applied could also be responsible for the different variances. It should also be noted that the instructions for the *z*-transformation

of the procedure's raw data are now 30 years old. Similarly, the subjects in our study had problems answering the questions. As a result, it was necessary to correct a number of false cumulative values that arose during the procedure, which involved distributing 15 points over 8 possible answers to a question.

In setting up the System I (intuitive) and System II (sensing) variables, the individual results of the NEO-FFI test to identify the Big Five personality traits were collected in accordance with Furnham's theories, as described earlier. This resulted in a four-step scale ranging from 1 (very slight characteristics) to 4 (very strong characteristics). In line with Kahneman's theories, the two systems were considered as being independent and the results are reported separately. Both systems produced on average figures in the average range (2.64/2.35), but the "intuition" variable was more widely distributed.

Table 3. descriptive statistics of current study contd. (age, IQ, AD stage, intuitive and sensing)

	age	CFT IQ	ego development TPR	gravity levels of existence	system 1 (intuitive)	system 2 (sensing)
mean/median	42.26	123.78	self-aware (3/4)	relativistic (F−S)	2.64	2.35
SD	10.23	11.80	—	—	0.43	0.29
min	18	94	conformist (3)	reactive (A−N)	1.56	1.61
max	61	153	autonomous (5)	experiantialistic (H−U)	3.53	3.11

Testing of hypotheses:

Hypothesis 1a: Managers show a higher degree of development with regard to adult development theories than non-managers.

In order to test this hypothesis, a *t*-test was run using managers as the grouping variable and two test variables:

a) Centre of gravity for the Levels of Existence

b) Total protocol rating in the Sentence Completion Test

In the Sentence Completion Test there was a highly significant correlation between group membership and rating. All six high-level protocols were also attributed to the management group sample. This correlation could not be established for the centre of gravity for the Levels of Existence. So the results are varied, with the hypothesis for Ego Development being proven and the hypothesis for the Levels of Existence being rejected. Managers differ from our highly selective sample of non-managers in their level of Ego Development, in the sense of a higher level of development, but not in the Levels of Existence.

Hypothesis 1b: Managers make abstract value choices during execution of the FMRI paradigm that demonstrate a tendency towards the values of a higher Level of Existence.

Table 4. independent samples *t*-test comparing stage of ego-development, levels of existence and choices of words related to the levels of existence depending on leadership role

variable	*F*	*p*
ego development	12.48	< 0.01
gravity levels of existence	0.46	ns
neurochoices	1.98	ns

Table 5. independent samples *t*-test comparing intuition and sensing personality patterns depending on leadership role

variable	F	p
system I (intuition)	1.38	ns
system II (sensing)	1.64	ns

To test this hypothesis, a *t*-test with leadership as grouping variable and "neurochoices" as test variable was carried out, producing a weighted score for the choices made in the FMRI scenario. This added up all the choices made for the individual Levels of Existence and calculated an overall total from the cumulative values using a simple ascending weighting factor for the individual stages.

It was also not possible to establish a statistically significant correlation to membership of one of the two groups (managers/non-managers) (cf. Table 4). Therefore hypothesis 1b has to be rejected.

Hypothesis 2a: Managers display a personality type that is closer to the "intuitive" type than non-managers.

In order to test this hypothesis, a *t*-test was carried out using leadership as grouping variable and System I (intuitive) as test variable. The variable corresponded to Furnham's theories, as described earlier.

Hypothesis 2b: Non-managers display a personality type that is closer to the "globally perceptive" type than managers.

In order to test this hypothesis, a *t*-test was carried out using leadership as grouping variable and System II (globally perceptive) as the test variable. The variable corresponded to Furnham's theories, as described earlier.

It was not possible to establish a statistically significant correlation to membership of one of the two groups (managers/non-managers) for either the "intuitive" or the "globally perceptive" personality traits. Further *t*-tests with the five NEO-FFI factors only produced a negative correlation between neuroticism and the management role ($p < 0.05$) and a positive correlation between agreeableness and the management role ($p < 0.05$). Therefore hypothesis 2 (a and b) has to be rejected.

Hypothesis 3a: The "intuitive" personality type correlates strongly to a higher level of adult development than the "globally perceptive" personality type.

In order to test this hypothesis, a multiple linear regression was used with the following independent variables:

1. Total protocol rating in the Sentence Completion Test
2. Centre of gravity for the Levels of Existence and the predictors system I (intuitive) and system II (globally perceptive).

In the multiple linear regressions, only the "intuitive" predictor variable makes a minor contribution to explaining the total variances of ego development ($R = 0.40$, corresponding to 16% of the total variance) and levels of existence ($R = 0.33$, corresponding to 11% of the total variance). Hypothesis 3a can therefore be considered as proven.

Table 7. effects of personality clusters "intuition" and "sensing" on center of gravity in values test

intuition	levels of existence	sensing
$ß = 0.33$	$R = 0.33$	$ß = -0.08$
$p < 0.01$	$R^2 = 0.11$	$p = $ ns

Hypothesis 3b: The "intuitive" personality type is linked to choices made during execution of the FMRI paradigm that can be attributed to higher levels of existence.

In the multiple linear regressions, only the "intuitive" predictor variable makes a minor contribution to explaining the total variances of choices made during execution of the FMRI paradigm ($R = 0.21$, corresponding to 6% of the total variance). Hypothesis 3b can therefore be considered as proven.

Hypothesis 4a: The "globally perceptive" personality type is linked to average levels of adult development.

Hypothesis 4b: The "globally perceptive" personality type is linked to a tendency to make choices during execution of the FMRI paradigm that can be attributed to the middle levels of existence.

In order to test hypotheses 4a and 4b, the aforementioned multiple linear regressions were enhanced by individual correlations of the two personality types with the various overall measures of adult development and particular characteristics at specific stages in order to attempt to find evidence of U-shaped relationships.

In this respect it was generally assumed that attempts to perceive situations in a global way are more often displayed in the middle stages of adult development. This is because in the lower stages there is an aversive rejection of complexity and in the higher stages there is a strongly intuitive recognition of patterns that tends to replace global perceptiveness.

Whereas for the "intuitive" type the pattern of correlations that is consistent with the hypothesis shows significant negative correlations at lower levels and significant positive correlations at higher levels. This was true for both Ego Development and the Levels of Existence. For the "perceptive type" the picture was less uniform. A significant correlation with the "perceptive type" can only be identified in the middle stage of ego development, which corresponds with theoretical expectations. But for the levels of existence we see a picture that is the exact opposite. Here the earliest B–O level in our study has a negative correlation to the "perceptive type" and the highest level has a positive correlation to the "perceptive type". As far as the choices during execution of the FMRI paradigm are concerned, there was no significant correlation to the "intuitive" type and no correlation to the "perceptive" type. However, there was a positive correlation to the total value and a negative correlation to the choices made at the earliest level studied,

Overall, hypothesis 4a can only be proven for Ego Development. For the Levels of Existence, the results can be explained by the fact that the B–O level is generally less concerned with or capable of

Table 6. effects of personality clusters "intuition" and "sensing" on WUSCT total protocol rating

intuition	ego development	sensing
$ß = 0.41$	$R = 0.40$	$ß = -0.04$
$p < 0.01$	$R^2 = 0.16$	$p = $ ns

Table 8. effects of personality clusters "intuition" and "sensing" on neurochoices

intuition	neurochoices	sensing
$ß = 0.19$	$R = 0.21$	$ß = -0.06$
$p < 0.05$	$R^2 = 0.06$	$p = $ ns

Table 9. correlations of personality clusters "intuition" and "sensing" with adult developmental measures

variable	intuition	sensing
ego development	0.39‡	0.09
no. conformist SC	−0.22†	−0.01
no. self-aware SC	−0.27‡	−0.08
no. conscientious SC	0.41‡	0.21†
no. higher SC	0.22†	−0.10
levels of existence	0.33‡	0.11
z-value B−O	−0.31‡	−0.26‡
z-value C−P	−0.11	−0.14
z-value D−Q	−0.36‡	0.01
z-value E−R	0.09	0.12
z-value F−S	−0.17	−0.16
z-value G−T	0.32‡	0.25‡
neurochoices	0.20†	0.12
choose B−O	−0.24‡	−0.02
choose C−P	0.14	−0.17
choose D−Q	−0.16	0.09
choose E−R	0.15	0.15
choose F−S	−0.09	−0.12
choose G−T	0.17	0.10

† $p < 0.05$; ‡ $p < 0.01$

individual data collection and complex processing of information. In contrast, the F−S level of existence prefers this complexity and immediately addresses new information.

Hypothesis 5a: The "neuroticism" personality trait is linked to a low level of adult development.

Hypothesis 5b: The "neuroticism" personality trait is linked to a tendency to make choices during execution of the FMRI paradigm that can be attributed to the lower levels of existence.

A one-way ANOVA with the "neuroticism" factor (z-transformed values) was used to test this hypothesis, along with the following three dependent variables:

1. Total protocol rating in the Sentence Completion Test

2. Centre of gravity for the Levels of Existence

3. Choices made during execution of the FMRI paradigm.

As expected, there was a significant negative correlation between the state of development in the Levels of Existence and neuroticism. However, this correlation could not be proven for Ego Development. Therefore hypothesis 5a could only be partially proven. It is likely that the Levels of Existence are more strongly bound up with emotions—particularly the degree of aversive activation of the circumplex model of emotion—while ego development is more closely linked to the System I and System II activation. However, this should be qualified by noting that this correlation to the forced choice scenario as implemented in the FMRI paradigm could not be replicated.

» GENERAL DISCUSSION

All in all, this study was able to throw an exploratory light on the correlation between different adult development theories and different types of personalities and emotions on the background of leadership responsibility. The notional "intuitive" and "globally perceptive" personality types were replicated and linked to the status of Ego Development. In contrast, the measurement of Levels of Existence seems to have a greater link to a negative activation component, as described in the circumplex models of emotion. So for the first time this indicates a potentially comprehensive integration of neuroscientific and psychological theories and findings with the three activation components of negative affect, intuitive processing and cognitive appraisal. These are correlated to the corresponding neurological processes and also to the corresponding stages in adult development.

Of course this study also shows many gaps that could be filled by further investigations in this area of research. The questionnaire-based processes used are out of date and were designed for completely different groups of subjects (girls and women in general for the WUSCT, rather than leaders). Too few statistical validation studies have been carried out ending up in insufficient results concerning test reliability. Similarly, the structure of the questionnaire used proved to be too complicated for the test subjects. Data collection of values furthermore has some special restrictions regarding answer tendencies of test persons especially in the field of an assumed social desirability. From a research economical perspective, numerous questions arise about such a comprehensive study. The time required to carry out the tests, particularly the WUSCT, was criticized by many of the participants, which may to some extent have contributed to the shortness of the answers given, making them difficult to evaluate during the rating process.

With a view to future research, it would be a useful next step to use FMRI data to gain further insights into the neurobiological correlates of the postulated System I and System II variables created using Furnham's personality clusters.

It would also be desirable to use a longitudinal design for future questionnaire-based research. This is because it is only possible to make hypothetical statements about the causal relationships of the data studied when it is based on a cross-sectional design. It would therefore be very useful to repeat this kind of study at various times on the same subjects. The first author is currently developing a free web-based test system for researchers in order to provide a comprehensive and yet research-economic means of carrying out longitudinal research in this area. Please contact us for more details. ∎

Table 10. one-way ANOVA comparing levels of existence, stage of ego-development and choices of words related to the levels of existence depending on neuroticism

variable	F	p
ego development	1.17	ns
gravity levels of existence	2.00	< 0.01
neurochoices	1.29	ns

REFERENCES

Aellig, S. (2004). Über den Sinn des Unsinns: Flow-Erleben und Wohlbefinden als Anreize für autotelische Tätigkeiten, Waxmann Verlag, Münster.

Beer, S. (1995). *Diagnosing the System of Organizations*. Chichester: Wiley.

Beck, D. & Cowan C. (2007). Spiral Dynamics - Leadership, Werte und Wandel: Eine Landkarte für das Business, Politik und Gesellschaft im 21. Jahrhundert. Kamphausen, Bielefeld.

Brown, B. C. (2012). Leading complex change with post-conventional consciousness. *Journal of Organizational Change Management*, *25*(4), 560–577.

Camerer, E., Loewenstein, G. & Prelec, D. (2005): Neuroeconomics – How Neurosciences can inform economics. *Journal of Economic Literature*, *43*, 9–64.

Caspers, S., Heim, S., Lucas, M. G., Stephan, E., Fischer, L., Amunts, K., & Zilles, K. (2011). Moral Concepts Set Decision Strategies to Abstract Values. *PLOS One*, *6*(4), e18451. Doi: 10.1371/journal.pone.0018451.

Caspers, S., Heim, S., Lucas, M. G., Stephan, E., Fischer, L., Amunts, K. & Zilles, K. (2012). Dissociated neural processing for decisions in managers and non-managers. *PLOS One*, *7*(8), e43537. Doi: 10.1371/journal.pone.0043537.

Commons, M. L., Armon, C., Richards, F. A., Schrader, D. E., Farrell, E. W., Tappan, M. B., et al. (1989). A multidomain study of adult development. In D. Sinnott, F. A. Richards & C. Armon (Eds.), *Adult development, Vol. 1: Comparisons and applications of developmental models* (pp. 33–56). New York: Praeger Publishers.

Commons, M. L., Goodheart, E., Bresette, L., Bauer, N. F., Farrell, E. W. & McCarthy, K. G. (1995). Formal, systematic and metasystematic operations with a balance-beam task series: A reply to Kallio's claim of no distinct systematic stage. *Journal of Adult Development*, *2*(3) 193–199.

Cooper, P. (2005). Developing transformational leadership capacity in the public service. *Public Administration Today*, *3*, 66–76.

Corbett, R.P. (1995): *Managerial style as a function of adult development stage*. Unpublished Dissertation, University of Massachusetts, Amherst.

Fischer, L. & Wiswede, G. (2009). Grundlagen der Sozialpsychologie, Oldenbourg Wissenschaftsverlag, München.

Fischer, R. A. (1971). A cartography of the ecstatic and meditative states. *Science*, *174*(26), 132–146.

Furnham, A. (1996). The Big Five versus the Big Four: The relationship between the Myers-Briggs Type Indicator (MBTI) and NEO-PI Five factor model of personality. *Personality and Individual Differences*, *21*(2), 303–307.

Furnham, A. & Stringfield, P. (1992). Personality and work performance: Myers-Briggs Type indicator correlates of managerial performance in two cultures. *Personality and Individual Differences*, *14*(1), 145–153.

Graves, C. W., Huntley, W. C. & LaBier, D.W. (1965). Personality structure and perceptual readiness: an investigation on their relationship to hypothetisized levels of human existence. Retrieved September, 2014, from www.claregraves.com/articles.html.

Graves, C.W. (1970). Levels of existence: An open system theory of values. *Journal of Humanistic Psychology*, *10*, 131–155.

Graves, C. W. (1974). Human nature prepares for a momentous leap. *The Futurist*, *4*, 72–87.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M. & Cohen, J. D. (2004). The neural basis of cognitive conflict and control in moral judgment. *Neuron*, *44*(2), 389–400.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M. & Cohen, J. D. (2001). A fMRI investigation of emotional engagement in moral judgment. *Science*, *293*, 2105–2108.

Gurtman, M. B. (1997). Studying Personality Traits: The Circular Way. In: Robert Plutchik & Hope R. Conte (Eds.): *Circumplex models of personality and emotions*, (pp. 81–102). Washington: APA.

Haidt, J. (2001). The emotional dog and it's rational tail: a social intuitionist approach to moral judgement. *Psychological Review*, *108*, 814–834

Hamilton, M. (2012). Leadership to the power of 8: Leading self, others, organization, system and supra-system. *Wirtschaftspsychologie*, *14*(3), 58–64.

Hy, L. X. & Loevinger, J. (1996). *Measuring Ego Development*. (2nd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Ilies, R., Gerhardt, M. & Le, H. (2004). Individual differences in leadership emergence: Integrating meatanalytic findings and behavioral genetics estimates. *International Journal of Selection and Assessment*, *12*(3), 207–219.

Jacques, E. (1998). *Requisite Organizations*. Arlington: Cason Hall & Co.

Judge, T. A., Bono, J. E., Ilies, R. & Gerhardt, M. W. (2002). Personality and leadership. *Journal of Applied Psychology*, *87*(4), 765–780.

Kahneman, D. (2003). Maps of bounded rationality: A perspective on intuitive judgment and choice. *American Economic Review*, *93*(5), 1449–1475.

Kahneman, D. (2011). *Thinking Fast and Slow*. Farrar, Strauss and Giroux: New York.

Kallio, E. & Pirttilä-Backman, A.-M. (2003). Developmental processes in adulthood – European research perspectives. *Journal of Adult Development*, *10*(3), 135–138.

Karsten, E. (2006): *Spiral Dynamics onderzocht – Een studie naar de validering en betrouwbaarheid van vragenlijsten op het gebied van spiral dynamics*. Unpublished Master Thesis, University of Amsterdam.

Kohlberg, L. (1969). State and sequence: The cognitive-developmental approach to socialisation. In: D. A. Goslin (Ed.): Handbook of socialization theory and research, 347–480, Rand McNally: Chicago.

Küppers, B. O. (2000). Die Strukturwissenschaften als Bindeglied zwischen Natur- und Geisteswissenschaften. In B. O. Küppers (Ed.): *Die Einheit der Wirklichkeit*, (pp. 89–105). München: Fink.

Loevinger, J. (1970). *Measuring Ego Development*. Jossey-Bass: San Francisco.

Loevinger, J. (1976). Ego development. Conceptions and theories. Jossey-Bass: San Francisco.

Loevinger, J. (1993). Ego development: Questions of method and theory. *Psychological Inquiry*, *4*, 56–63.

Lucas, M. (2012). Grundlagen einer neo-integralen transformationalen führung und organisationsentwicklung. *Wirtschaftspsychologie*, *3*, 4–20.

Lucas, M. (2013). Foundations of a neo-integral transformational leadership and organizational development. *Integral Leadership Review*, *1*, 12–27.

Marchand, H. (2001). Some reflections on postformal thought. *The Genetic Epistemologist*, *29*, 3.

Merron, K. (1985). *The relationship between ego-development and managerial effectiveness under conditions of high uncertainty*. Unpublished Dissertation, Harvard University

Peters, T. & Ghadiri, A. (2011). *Neuroleadership – Grundlagen, Konzepte, Beispiele*. Gabler: Wiesbaden.

Plutchik, R. (1997). The Circumplex as a General Model oft he Structure of Emotions and Personality. In: Robert Plutchik & Hope R. Conte (Eds.): *Circumplex models of personality and emotions*, (pp. 17–46). Washington: APA.

Quinn, R. & Torbert, W.R. (1987). *Who is an effective transforming leader?* Unpublished paper, University of Michigan.

Robinson, O. (2013): *Development through Adulthood*. Palgrave: Basingstoke.

Rooke, D. & Torbert, W. R. (1998). Organizational transformation as a function of CEO's developmental stage. *Organizational Development Journal*, *16*(1), 11–28.

Schallberger, U. (2006). Die zwei gesichter der arbeit und ihre rolle für das wohlbefinden: Eine aktivierungstheoretische interpretation. *Wirtschaftspsychologie*, *2*(3), 96–102.

Schallberger, U. & Pfister, R. (2001). Flow-Erleben in arbeit und freizeit. Eine untersuchung zum paradoxon der arbeit mit der Experience Sampling Method (ESM). *Zeitschrift für Arbeits- und Organisationspsychologie*, *45*, 176–187.

Schilke, O. & Reimann, M. (2007). Neuroökonomie: Grundverständnis, Methoden und betriebswirtschaftliche Anwendungsfelder. *Journal für Betriebswirtschaft*, *57*, 247–262.

Schwartz, S. H. & Bardi, A. (2001). Value hierarchies across cultures: Taking a similarities perspective. *Journal of Cross-Cultural Psychology*, *32*(3), 268–290.

Schwartz, S. H. (2006): Basic human values: Theory, measurement, and applications. *Revue francaise de sociologie*, *47*(4), 1–43.

Smith, S.E. (1980). *Ego development and the problems of power and agreement in organizations*. Unpublished Dissertation, Fielding Institute.

Strack, M. (2011). Graves werte-system, der wertekreis nach schwartz und das competing values modell nach quinn sind wohl deckungsgleich. *Wissenswert*, *1*, 18–24.

Watson, D. & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, *98*(2), 219–235.

Watson, D. et al. (1999). The two general activation systems of affect. *Journal of Personality and Social Psychology*, *76*(5), 820–838.

Weibler, J. (2012). Personalführung, Vahlen, München

Goal pursuit and eudaimonic well-being among university students: Metacognition as the mediator

Yalda Amir Kiaei and Thomas G. Reio, Jr.

Florida International University

ABSTRACT

This study investigated the relationship between goal-striving, goal-aspiration, metacognition, and eudaimonic well-being (EWB). Inspired by Aristotle's teaching, the rationale for this study is that eudaimonic well-being is achievable through self-actualizing processes such as goal-striving and goal-aspiration and by exercise of reason. Goal-striving, metacognition (as a way of exercise of reason), and goal-aspiration (as an indicator of eudaimonic pursuits) were explored in relation to EWB. A mediation analysis of a sample of 513 university students (M_{age} = 25.07, SD = 7.21) indicated that metacognition partially mediated the relationship between goal-striving and EWB for the full sample ($p < .001$) and goal-aspiration moderated this relationship. High goal-aspiration indicated a full mediation while low goal-aspiration indicated only a partial mediation. The finding suggests that metacognition which is a teachable competence and goal-aspiration which is a trainable desire can play a determining role in individuals' self-actualization and EWB.

KEYWORDS: eudaimonic well-being, self-actualization, goal-striving, goal-aspiration, well-being, metacognition, emerging adults, eudaimonia, happiness, goal pursuit, goal orientation

WELL-BEING AND HAPPINESS have been of great import for human beings since ancient times. Scholarly efforts to shed light on these concepts can be traced to Hellenic philosophy when the closely linked concept of *human goods* was of concern. Yet, the route towards a fulfilling life remains complicated and ambiguous; so much so that it may never be clarified entirely. Nevertheless, research conducted with untested new variables could augment our understandings of the path to well-being and happiness. Therefore, the present study investigated some relatively untested possible elements that may help to pave the route to, what may be considered an individual's well-being.

As such, the purpose of this study was to investigate the relationship between goal-pursuit (as a self-actualizing process), metacognition (as an attainable intellectual capacity), and (eudaimonic) well-being. In this attempt, the present study was inspired by Aristotle's teaching in *Nicomachean Ethics* about the human ultimate good and the requirements of living a good life. The concepts of self-actualization and metacognition, as they will be discussed in the conceptual framework, address elements of character and intellect in Aristotle's conception of the human good to some levels. The concept of eudaimonic well-being addresses the concept of a good life at a secondary level and in a personal and subjective sense.

Author note: correspondence regarding this article should be directed to Yalda Amir Kiaei, Florida International University, yamir001@fiu.edu.

» THE CONCEPT AND CONCEPTIONS OF HUMAN WELL-BEING

Some of the contemporary conceptual works on human well-being (Ryan et al., 2008; Waterman, 1990A, 1990B) made references to *eudaimonia*. Eudaimonia is a Greek term, which is mainly known by the special attention it received in *Nicomachean Ethics*. Nicomachean Ethics is a collection of lectures of Aristotle, the philosopher of the fourth century BCE The concern with eudaimonism and human goods, however, predates Aristotle (Norton, 1976). The concept of eudaimonia or well-being, in Greek philosophy, is commonly understood as the ultimate good and the purpose of life for human beings. Yet as Aristotle (2004) also acknowledged, people used (and are using) the same term eudaimonia (well-being or happiness) with different conceptions in mind. Thus, ancient and contemporary philosophers are also not exempt from representing varying conceptions of the ultimate good.

For instance, there is a conception of the ultimate good, pioneered by Aristippus in the third century BCE, that favors pleasure and feeling good (i.e., hedonic enjoyment) over any other goods (Tatarkiewicz, 1976). In the modern literature, there is another conception of well-being, namely, hedonic or subjective well-being. It is different from hedonic enjoyment or pleasure in that its concept commonly includes the cognitive judgment of satisfaction with life in addition to positive affect and negative affect (Andrews & Withey, 1976; Deci & Ryan, 2008; Diener, 1984; Diener, Suh, Lucas, & Smith, 1999).

On the other hand, Aristotle (2004) gave a very objective account of human well-being arguing that eudaimonia is the highest, the noblest, and the most complete of all goods, which is a contemplative activity of the soul. Because he was concerned with the good of the society, he noted that contemplative work, the greatest good for an individual human, needed to be manifested in practice if it was to influence the society. Therefore, his work was more focused on being a *good human,* who also has a happy well-lived life as a secondary eudaimonic life compared to a wise human. This secondary eudaimonic life, he argued, was at its best the life of a practical wise person who lived and acted in accordance with both virtues of character (e.g., courage, honesty) and intellect (i.e., judgment, wisdom). He also argued that virtues of intellect are achievable mainly through education and exercise of reason, and virtues of character were attainable mainly by habituation and training. This quality, of living well and acting well, needed to be deliberately pursued and desired and chosen for its own sake and for nothing else.

Moreover, there is another conception of eudaimonia that focuses on *daimon* as the true self and interprets eu-daimon-ia (*eu* means well) as living in accordance with the true self. This translates *o* knowing and living in truth to one's true self through actualizing one's own innate strength or potential excellence (Norton, 1976). This is the conception that has driven contemporary eudaimonic approaches in the conceptualization of human well-being (e.g., Ryan, Huta, & Deci, 2008; Seligman, 2002; Waterman, 1990A). Waterman (1990A, 1990B) was primarily guided by Norton's (1976) contemporary representation of eudaimonism and influenced by the work of other contemporary philosophers (e.g., Kraut, 1979; Telfer, 1980). He presented the concept of *eudaimonic well-being* as an individualistic and subjective account of eudaimonia. He argued that this subjective account of eudaimonia was more practical in the realm of psychological science than that of Aristotle. He, in line with Norton (1976), argued that the concept of ultimate good involved actualizing those potentialities which represent the true self (i.e., the best that one is capable of being) unique to each individual person.

Conceptual framework

Being concerned with individuals' well-being in terms of being-the-best-one-can-be, the authors of the present study are interested in the human good that involves one striving for his/her own best that s/he is capable of being. This conception of well-being, which is more in line with Waterman's (1990A, 1993), involves actualizing one's capabilities, capacities, and competences in talent-related, skill-related, character-related, and intellectual pursuits and any other growth-related goal-oriented efforts. This conception of well-being suggests that well-being represents a life which involves self-actualization as

> *ongoing actualization of potentials, capacities, and talents, as fulfillment of mission (or call, fate, destiny, or vocation), as a fuller knowledge of , and acceptance of, the person's own intrinsic nature, as an increasing trend toward unity, integration, or synergy within the person. (Maslow, 1968, p. 25)*

A review of related conceptual literature (e.g., Maslow, 1968; Rogers, 1951, 1961; Ryan et al., 2008; Seligman, 2002; Waterman, 1990A, 1990B) also suggested that a conceptual link between well-being and self-actualizing efforts was implied by the literature. However, the nature of this relationship and, in fact, whether or not they are two distinct phenomena, is rather ambiguous in the existing literature. Some assistance from Aristotle's teaching may shed light on this issue.

Although this study is concerned with eudaimonic well-being, which is significantly different from Aristotle's eudaimonia in several ways (see Waterman 1990A for detailed discussion), the authors tried to have considerations of the Aristotelean eudaimonic view in their conceptualization and operational definitions of the concepts that are hypothesized as antecedents to eudaimonic well-being. The gist of Aristotle's (2004) teaching, as evidently presented in the last three chapters of *Nicomachean Ethics*, is that achieving excellence in intellect is superior to other kinds of achievement in terms of what is conceptually the best in and for humans. But, again, he argued that practically what is best for us is excellence in both character and intellect; he stated that this practically good life is happy/well-lived, but it represents eudaimonia in a secondary way.

We may logically argue that these concluding remarks from Aristotle implies that any kind of *human* characteristic activity contributes to one's well-being at some level (not the highest level though) depending on the level of character and intellect at which the respective characteristic activity is being performed. Following the same line of argument as Aristotle's (2004), only at a subjective and secondary level however, the less ideal yet fulfilling life (i.e., eudaimonic well-being) may be attained if we live and act in accordance with the bests of our personal goals through exercise of reason and rational choice. These personal goals may be talent-related, skill-related, health-related, relationship-related, and so on. We argue that eudaimonia might be achieved by living and acting in accordance with virtue of character and intellect through exercise of "goal-directed reason" (representing intellect; p. 104) and rational choice (representing character and intellect; p.104). We also argue that any of these can be accomplished at differing levels. With different levels of (right) rational choice, we will have different levels of goal-aspiration. Similarly, with different levels of exercise of reason, we will have different levels of performance in our goal-striving, and thus we will have different levels of eudaimonic well-being.

Relying on the above argument, the conceptual framework set forth in the present study conceptualizes self-actualization as a goal-oriented process. This goal-oriented process involves both goal-aspiration and goal-striving and eudaimonic well-being as a quality of life which is a concurrent outcome of this process. Further, we argue that metacognition, as an attainable intellectual capability and a way of exercise of goal-oriented reason, is essential for setting self-actualizing needs and goals and for facilitating the regulation of personal efforts towards those goals. That is, metacognition, which is primarily known as cognition about cognitive processes (Flavell, 1979; Lin & Zabrucky, 1998; McCromick, 2003), can serve the purpose of self-actualization by helping self-actualizing goal-setting and goal-striving. Thus, we also argue that, as

an overarching construct, metacognition can be hypothesized to mediate the relationship between self-actualizing variables (i.e., goal-aspiration and goal-striving) and eudaimonic well-being.

Eudaimonic well-being and self-actualization

Eudaimonic well-being has been operationally defined by Waterman et al. (2010) based on a contemporary account of eudaimonia. This account was significantly different from that of Aristotle's (see Waterman, 1990A). They introduced the concept of personal expressiveness as a quality that represents eudaimonic aspects of well-being and as a signifier of self-realization and optimal human functioning. Expressiveness occurs when individuals engage in actions which are in the direction of discovering and developing their potentials and oriented towards achieving their personal goals while adding meaning to their lives (Waterman et al. 2010). Thus, eudaimonic well-being is centered on personal expressiveness in the form of engagement and meaning in life. Engagement and meaning were also highlighted by Peterson, Park, and Seligman (2005) as two of the three routes to happiness (the third being pursuing pleasure [i.e., hedonic well-being]) and as eudaimonic aspects of a *full* life.

Eudaimonic well-being requires living an engaged and meaningful life. To acquire and maintain this lifestyle, human individuals need to be attuned to their inner self. They also need to continuously actualize their characteristic activities, the potentials actualizing which make them who they willingly wish to be. The term self-actualization seems to be the most plausible terminology to describe this fulfilling mindset and lifestyle. A review of literature on self-actualization indicated that the term self-actualization has been considered to imply similar connotations as do concepts such as maturation, full humanness, and ultimate being (Maslow, 1968) or full functionality (Rogers, 1951). However, research literature has loosely defined and used the term self-actualization. For instance, Kerr (1991) viewed self-actualization as *challenging the limits of intellectual potentials* to use the intellectual gifts to the fullest. He, however, reduced its operational definition to *receiving advanced academic degrees or prestigious jobs* (Kerr, 1985). Reis and his colleagues (Reis & Callahan, 1989; Walker, Reis & Leonard, 1992) considered it *the achievement in recognized fields of endeavor* such as in the areas of National Merit semi-finalists, patents, the United States Senate, and the Supreme Court.

Self-actualization: a goal-oriented process

Although the contemporary literature has not yet come to a consensus to operationalize the concept of (self-)actualization, Rule (1991) suggested a new approach to both conceptually and operationally defining the concept. Challenging the traditional view of self-actualization (e.g., Jones & Crandall, 1986; Shostrom, 1964) which was merely personality-oriented, Rule (1991) proposed self-actualization as a goal-oriented process. He argued that actualization denotes the process of growth, change, unfolding, developing, and transcending which manifests itself in personal goals people pursue (Rule, 1991, p.252). Human beings are growth oriented as much as they are goal-oriented. In fact, growth or self-actualization is the goal for the self (Rule, 1991). Rule's conceptualization of self-actualization implies that the quality of striving for personal goals

is a reflection of one's effort in the process of self-actualization. Accordingly, the present study considered striving for personal goals a self-actualizing process and hypothesized that personal goal-striving predicts eudaimonic well-being.

From Aristotle's (2004) teaching, the aim and the means of a good life need to be driven by the right reason. That is the aim and actions should be in accordance with the virtues of character and intellect. Similarly, the aim needs not be chosen for the sake of anything else but its own sake and because it is worth pursuing in itself and it is good by itself. It is also desired, pleasant, valued, and worthy of choice to a good person and to human beings in general. Thus, the self-actualizing goals also need to fulfill certain qualities to be eudaimonic. These qualities are also emphasized by contemporary eudaimonist scholars. Accordingly in this study, self-actualization was defined as pursuing and striving to achieve self-actualizing goals which are

a) inspired by eudaimonic and intrinsic aspirations (e.g., self-acceptance, relatedness, & helpfulness; Kasser & Ryan, 1996);

b) personally valuable and bring joy and pleasure to one rather than being imposed by other people or pursued to avoid negative feelings such as guilt or shame (i.e., self-concordant goals; Ryan & Connell, 1989; Sheldon, 2004);

c) pursued because they are good by themselves not because they are the means to achieve other goals (i.e., constitutive vs. instrumental goal orientations; Fowers, Mollica, & Procacci, 2010);

d) intrinsically worthwhile to human beings (i.e., virtuous; Ryan et al. 2008; Seligman, 2002); and

e) personally expressive (Waterman et al., 2010).

The variable goal-aspiration was conceptualized and measured to capture the range of these qualities. It was also investigated as a moderator in relation to goal-striving in predicting eudaimonic well-being.

Metacognition

There are some external and internal requirements for an effective goal pursuit to be taken place. The internal requirements may include emotional, cognitive or any other intrapersonal skills and the external requirements may take in any environmental factor from the physical environment to social interactions and skills. In this study, the authors focused on internal requirements for goal-striving and among them especially on cognition and metacognition. Metacognition, by definition, is *cognition about cognitive processes* (Flavell, 1979; Lin & Zabrucky, 1998; McCromick, 2003) and involves metacognitive knowledge and metacognitive regulation (Brown 1977, 1987). Metacognition is a trainable capacity and can be taught by teaching about how the brain works and as a set of regulatory skills and strategies (Borkowski & Muthukrishna, 1992; Gaskins & Pressley, 2007). The capacity of metacognition as a teachable and attainable skill for goal-setting and goal pursuit made it a proper candidate for a way of exercise of goal-oriented reason and, thus, to serve the purpose of this study.

In the metacognition framework, supplying the mind with personalized goals, tasks, and action plans could be referred to as *goal-setting* and *planning*. Goal-setting has been identified as one of the key functions of metacognition (Tarricone, 2011). Goal-setting, planning, and strategy-adjustment during a deliberate problem-solving situation, which equips people for their personal striving, are considered metacognitive regulatory skills (Brown, 1987; Tarricone, 2011) and involve both monitoring and control functions of the brain. Goal-Striving requires a combination of regulatory processes: *monitoring* the situation, identifying the problem or defining a personal project, setting goals and sub-goals to overcome the problem or accomplish the project, and *controlling* the actions or strategies being used. All these metacognitive regulatory skills in order to function require metacognitive knowledge including knowledge of strategies and knowledge of metacognitive processes involved. Metacognitive knowledge is a form of self-awareness (Brown, 1977; Brown & Smiley, 1977) which informs the regulatory processes (Tarricone, 2011).

Altogether, metacognition is characterized as an interacting combination of knowledge and skills, that involves monitoring and control processes during a deliberate problem-solving situation. In this study, metacognition was theoretically identified as a supporting factor for high quality personal striving and thus hypothesized to have a mediational effect in the relationship between goal-striving and well-being. The speculation is that metacognition could be a potential intrapersonal mediator for goal-striving in enhancing eudaimonic well-being.

For instance, assume that we have developing and maintaining a good relationship with our friends as a self-actualizing goal. Metacognition, as a way of exercise of reason, has been argued to help with identifying the problems or obstacles at the time or what can be improved. We can set the identified problems as sub-goals and tasks of metacognition and look for strategies that offer a solution to our problems (e.g., being more appreciative). Then, we try and switch from one strategy to another (e.g., being more kind when she needed you) if the first strategy is not working. One also has to be conscious and decisive about the goals and think about the next strategy to get closer to the goals (in this case, our friends). Thus, the goal of metacognition is to be mindful about our self-actualizing goals. Our eudaimonic well-being was hypothesized to be enhanced as much as we find our personal pursuits meaningful and engaging.

» RESEARCH QUESTION

This study explored the relationships among eudaimonic well-being, goal-striving and goal-aspirations (as two self-actualization variables), and metacognition. The main research question was: Does metacognition fully mediate the relationship between goal-striving and eudaimonic well-being? The secondary research question was: Does goal-aspiration moderate this mediational effect? To answer these questions the

following hypotheses were tested: *(1)* Goal-striving will significantly and positively predict eudaimonic well-being; *(2)* Goal-Striving will significantly and positively predict general metacognitive competence; *(3)* Metacognition will significantly and positively predict eudaimonic well-being; *(4)* Goal-striving will not significantly predict eudaimonic well-being over and above metacognition; *(5)* The indirect effect of goal-striving (through metacognition) on eudaimonic well-being will be significantly different from zero *(6)* Goal-aspiration will moderate the relationship between goal-striving and eudaimonic well-being, *(7)* Goal-aspiration will moderate the mediated relationship between goal-striving and eudaimonic well-being where metacognition is the mediator.

» METHOD

The present study was part of a larger research study exploring the nature and antecedents of eudaimonic well-being. For data collection, a battery of paper-and-pencil questionnaires were administered in a university setting using measures of well-being, general metacognitive competence, goal-striving, goal-aspiration, and demographic variables.

Participants

The participants were 513 undergraduate ($n = 370$) and graduate students ($n = 143$) ranging in age from 17 to 60 years-of-age ($M = 25.07$, $SD = 7.21$) from a public research university in South Florida, USA. Approximately 83% were emerging adults ($n = 428$, between 17 and 29 years of age; 30 years of age being the turning point towards more demographic stability; Arnett, 2000; Rindfuss, 1991) and 16.6% were adults ($n = 85$, at or above 30 years of age). Hispanics comprised 64.3% of the sample ($n = 330$); 64.3% was female ($n = 330$), and 17.3% of the participants were married ($n = 89$). Almost 60% of the sample consisted of education students; 39.2% were students representing a wide range of other majors (i.e., science, engineering, and international relations). Table 1 displays the demographic data.

Table 1. frequency table of demographic variables ($N = 513$)

variables	freq.	%
age		
emerging-adult (17–29)	428	83.4
adults (30–60)	85	16.6
gender		
male	183	35.7
female	330	64.3
marital status		
married	89	17.3
single	424	82.7
level of education		
graduate	143	27.9
undergraduate	370	72.1
major		
education	312	60.8
non-education	201	39.2
ethnicity		
hispanic	330	64.3
non-hispanic	183	35.7

Research measures

The measure of eudaimonic well-being (MEWB)
Eudaimonic Well-Being (EWB) was measured using a 20-item modified version of the Questionnaire for Eudaimonic Well-Being (QEWB; Waterman et al., 2010) scale and 7 items from the 28-item Life-Regard Index-Revised (LRI-R; Debats, 1998) on a 5-point Likert scale (0 being *Strongly Disagree* and 4 being *Strongly Agree*). Due to changes made to QEWB and addition of items from LRI-R to the measure, the preliminary validity of the new measure of eudaimonic well-being was tested using exploratory factor analysis (EFA). Items which did not load on any of the factors (i.e., Meaning and Pleasure of Engagement factors) by factor coefficients at or above .30 were not included in the measurement (Tabachnick & Fidell, 2007). From the total of 28 items only 21 were retained, which made up the

new Measure of Eudaimonic Well-Being (MEWB). Cronbach's alpha for the new, overall 21-item MEWB was .89 (for the two subscales: 14-item Meaning in Life α = .90, 7-item Pleasure of Engagement α = .70). The final EWB score was computed by summing the 21 item scores (M = 61.52, SD = 12.24).

The measure of general metacognitive competence (MGMC)
Metacognition (MC) was operationalized as general metacognitive competence using a modified version of the 52-item Metacognitive Awareness Inventory (MAI; Schraw & Dennison, 1994). For the purpose of this study, four items were removed. Wording of some items were revised to reflect *general* problem-solving situations rather than learning situations merely focused on studying for a class or test. Due to the modifications to the measure and based on the recommendation of the authors of the MAI, a preliminary validity analysis (EFA) was conducted, which yielded a new scale which we will call the Measure of General Metacognitive Competence (MGMC). As a result of exploratory factor analysis, 46 items were retained on this new measure using the cut-off point of ≥.30 for factor coefficients (26 items for Regulation of Cognition and 20 items for Knowledge of Cognition). Final MC scores were computed summing the item-ratings on the 46 items (M = 339.14, SD = 5.49). Cronbach's alpha for the overall measure (46 items) was .96 (for the subscales: Regulation of Cognition α = .94, and for Knowledge of Cognition α = .92). The items were rated on an 11-point Likert-Scale from 0 to 10 based on the extent to which they were true for each participant (0 being *not at all*, 5 being *half of the time*, and 10 being *all of the time*).

Measures for self-actualization variables
Self-actualization was measured in two ways; that is, goal striving and goal aspirations. First, participants were instructed to list five personal goals:

> *Please think of your personal projects you want to accomplish, your goals you are concerned about when planning for your future, and goals that inspire you in your everyday life. Personal goals might involve various life areas, as for example study, family, friends, your own personal growth, leisure time, health, jobs, housing conditions, etc. Focus on long-term goals (e.g., to improve the relationship with a friend) rather than on single behavioral acts or short-term pursuits (e.g., to buy a present for a friend next week).*

Then participants were asked to rate and answer questions for each goal about their goal-striving and aspirations on a 7-point Likert scale.
The measure of goal-striving (MGS). Goal-striving (GS) as an actualizing process was measured through self-rating of six items about

a) Importance (one item from Emmons's (1999) Striving Assessment Scale);

b) Effort (one item from Emmons's (1999) Striving Assessment Scale);

c) Clarity (two items from Robitschek's (1998) Personal Growth Initiative Scale); and

d) Inspiration (two items; Milyavskaya et al., 2012) in regard to each goal.

Participants rated the extent to which each of six statements described them from 1 to 7 (1 being *not at all*, 4 being *moderately*, and 7 being *very much*). Cronbach's alpha for the MGS was .91. Scores were summed on these six items for each goal and total scores for goal-striving were obtained averaging goal-striving scores across the five goals (M = 33.65, SD = 4.20).

The measure of goal-aspiration (MGA). To preliminarily reflect both Aristotle's eudaimonic view and contemporary views of eudaimonic living in personal goal pursuits, the authors used 15 items (partly adopted, partly developed) representing the eudaimonic values presented in the conceptual framework to measure goal aspiration (GA) Participants were asked, for instance, to what extent they pursued each goal: "*because it is good by itself*" (a constitutive orientation), "*because it makes you a better person*" (a subjective eudaimonic orientation), "*because it makes you look better among your friends, family, or other people in general*" (an extrinsic aspiration), "*because it makes you use and build upon your potentials*" (a personal expressiveness orientation), "*because it makes you learn new things about yourself and/or the world*" (an intellect orientation), " *because it makes other people's lives better*" (a character orientation), "*Because you endorse it freely and value it wholeheartedly*" (an intrinsic aspiration; as in Ryan and Connell's (1989) locus of causality). They rated the statements from 1 to 7 (1 being *not at all for this reason* and 7 being *very much for this reason*).

A preliminary validity study of this measure, as expected, indicated two factors: Extrinsic aspirations (6 items) and Actualizing Aspirations (9 items). Cronbach's alpha for the 15-item MGA was .70 (for Extrinsic Aspiration α = .88, and for Actualizing Aspirations α = .87). Items representing Extrinsic Aspiration (i.e., image, fame, popularity, and money) were reverse coded, so high scores indicated low extrinsic motivation and vice versa. Scores on all items were summed for each goal and were averaged across the five goals to obtain the overall GA score (M = 78.36, SD = 9.06).

» PROCEDURE

Two steps were taken to answer the research questions related to the mediation model: *(1)* Baron and Kenny's (1986) method; and *(2)* Shrout and Bolger's (2002) Bootstrap Procedure. The first four hypotheses in essence followed the four Baron and Kinney (1986) steps for testing mediation models, with the exception that the authors included control variables in testing each hypothesis. Thus, a General Linear Regression approach (GLM) was used. Shrout and Bolger's (2002) Bootstrap Procedure using structural equation modeling (SEM) was taken to test the mediation model and to specify the indirect effect of goal-striving on eudaimonic well-being. Shrout and Bolger's (2002) Bootstrap Procedure has no assumption of normality (Efron & Tibshirani, 1993; Mooney & Duval, 1993) and has been shown to yield higher statistical power for small or moderate sample size than Baron and Kenny's (1986) method (MacKinnon et al., 2002). GLM was also used to test Hypothesis 6. The moderated mediation in Hypothesis 7 was tested using a multigroup approach in SEM (see Ng et al.)

Table 2. zero-order correlation coefficients for well-being variables, goal-striving, and general metacognitive competence ($N = 513$)

variables	EWB	GS	GA	GMC
EWB	1.00	—	—	—
GS	0.35*	1.00	—	—
GA	0.36*	0.44*	1.00	—
MC	0.56*	0.36*	0.27*	1.00

Note. SWB = subjective well-being; EWB = eudaimonic well-being; GS = goal-striving; GA = goal-aspiration; GMC = general metacognitive competence. * $p < 0.001$, one-tailed.

Pre-analyses

Before proceeding with hypothesis testing, the preliminary correlational analysis showed that all variables were positively and significantly correlated as predicted (see Table 2), with correlation coefficients ranging from .27 to .56. The homogeneity of regression slopes test (McNeil, Newman, & Fraas, 2012) for the full sample also indicated that in predicting EWB, there was no interaction between Metacognition and Goal-Striving, $F(1, 509) = 1.036$, $B = −.002$, $\beta = −.413$, $p = .31$, Metacognition and Goal-Aspiration, $F(1, 509) = .771$, $B = .001$, $\beta = .364$, $p = .38$, and Goal-Aspiration and Goal-Striving, $F(1, 509) = 3.229$, $B = −.021$, $\beta = −.908$, $p = .07$.

Furthermore, to be able to account for demographic variable effects in the subsequent analyses, the possible differentiating demographic effects on the research variables were examined using the GLM approach taking all the demographic variables and their two-way interactions into account. Students' age group, major, and gender were detected to have differentiating effect on variables of interest. However, no interaction effect was detected. Emerging adults scored significantly lower on eudaimonic well-being and metacognition compared to adults, respectively $B = −5.28$, $\beta = −.16$, $p < .001$ and $B = −20.76$, $\beta = −.14$, $p < .05$. The mean EWB score was 66.92 ($SD = 10.98$) for adults and 60.45 ($SD = 12.21$) for emerging adults; the mean GMC score was 356.46 ($SD = 51.51$) for adults and 335.70 ($SD = 55.67$) for emerging adults. In general, education students scored significantly higher on eudaimonic well-being, goal-striving, and goal-aspiration as compared to non-education students, respectively $B = 4.37$, $\beta = .17$, $p < .001$, $B = 1.55$, $\beta = .18$, $p < .001$, and $B = 2.29$, $\beta = .12$, $p < .01$. The mean EWB score was 63.56 ($SD = 12.06$) for education students and 58.36 ($SD = 11.88$) for non-education students. The mean GS score was 34.41 ($SD = 3.89$) for education students and 32.48 ($SD = 4.39$) for non-education students. The mean GA score was 79.57 ($SD = 9.16$) for education students and 76.47 ($SD = 8.57$) for non-education students. Further, males scored significantly lower on goal-striving and goal-aspiration compared to females, respectively $B = −1.05$, $\beta = −.12$, $p < .05$ ($M = 32.61$, $SD = 4.59$ for males and $M = 34.23$, $SD = 3.85$ for females) and $B = −2.29$, $\beta = −.12$, $p < .01$ ($M = 76.34$, $SD = 9.46$ for males and $M = 79.48$, $SD = 8.64$ for females). The final models that included the important demographic variables are presented in Table 3.

Based on these findings, in addition to testing the hypothesis on the full sample, the sample was split into two independent sub-samples by each respective demographic variable (i.e., gender, major, and age group); that is, three sets of two independent sub-samples, and the hypotheses were tested on each sub-sample to examine if we could find support for each hypothesis despite demographic differences. While dividing the sample by one demographic variable, the remaining demographic variables were used as control variables in testing the hypotheses as appropriate. Table 4 displays means and standard deviations for the variables of interest in the full sample and within each sub-sample. However, being mindful of space considerations, the results of the sub-sample analyses were reported only if they deviated from that of the full sample analyses.

Hypotheses testing and results

Hypothesis 1

Hypothesis 1 states that Goal-Striving positively and significantly predicts EWB. In the full sample, the GLM approach indicated that Goal-Striving explained 10.1% of variance in EWB, $F(1, 509) = 61.887$, $B = .95$, $\beta = .33$, $p < .001$, over and above age group and major of study which together accounted for 6.8% of variance in EWB. The regression coefficient also indicated that for every 1 SD change in Goal-Striving, EWB changes by .33 SD. Further, the analyses for the sub-samples also yielded similar results in support of Hypothesis 1. The unique contribution of goal-striving to eudaimonic well-being was the smallest for adults (5.6% of the variance) and was the greatest for males (14.1% of the variance). Overall, Hypothesis 1 was supported for all sub-samples as well as the full sample indicating a significant positive relationship between goal-striving and eudaimonic well-being.

Table 3. summary of GLM predicting eudaimonic well-being, metacognition, goal-striving, and goal-aspiration from demographic variables ($N = 513$)

variables	eudaimonic well-being						
	B	β	R^2	$\bar{R}^2$	ΔF	df_1	df_2
age group	−5.280†	−0.161	—	—	—	—	—
major	4.368†	0.174	—	—	—	—	—
overall model	—	—	0.068	0.064	18.538†	2	510
	metacognition						
	B	β	R^2	$\bar{R}^2$	ΔF	df_1	df_2
age group	−20.759*	−0.139	—	—	—	—	—
overall model	—	—	0.019	0.017	10.100*	1	511
	goal-striving						
	B	β	R^2	$\bar{R}^2$	ΔF	df_1	df_2
gender	−1.049*	−0.120	—	—	—	—	—
major	1.549†	0.180	—	—	—	—	—
overall model	—	—	0.063	0.059	17.006†	2	510
	goal-aspiration						
	B	β	R^2	$\bar{R}^2$	ΔF	df_1	df_2
gender	−2.293*	−0.121	—	—	—	—	—
major	2.286*	0.123	—	—	—	—	—
overall model	—	—	0.041	0.037	10.838†	2	510

Note. $\bar{R}^2$ = adjusted R-squared; ΔF = change in F. * $p < 0.01$; † $p < 0.001$.

Hypothesis 2

Hypothesis 2 states that Goal-Striving positively and significantly predicts metacognition. In the full sample, the GLM approach indicated that Goal-Striving explained 13.1% of variance in Metacognition, $F(1, 510) = 78.43$, $B = 4.78$, $\beta = .36$, $p < .001$, over and above age group which accounted for 1.9% of variance in MC. The regression coefficient also indicated that for every 1 SD change in Goal-Striving, MC changes by .36 SD. Further, the analyses for the sub-samples also yielded similar results in support of Hypothesis 2. The contribution of goal-striving to metacognition was the greatest for males (19.8% of the variance) and was the smallest for females as well as education students (10% of the variance). Overall, Hypothesis 2 was supported for all sub-samples as well as the full sample indicating a significant positive relationship between goal-striving and Metacognition.

Hypothesis 3

Hypothesis 3 states that metacognition positively and significantly predicts eudaimonic well-being. In the full sample, the GLM approach indicated that metacognition explained 28.1% of variance in eudaimonic well-being, $F(1, 509) = 219.63$, $B = .118$, $\beta = .54$, $p < .001$, over and above age group and education which together accounted for 6.8% of variance in EWB. The regression coefficient also indicated that for every 1 SD change in MC, EWB changes by .54 SD. In addition, the analyses for the sub-samples also yielded similar results in support of Hypothesis 3. The unique contribution of metacognition to EWB was the greatest for non-education students (31% of the variance) and closely followed by males (30.8%) and was the smallest for females (27.3% of variance) as well as education students (27.6% of the variance). Overall, Hypothesis 3 was supported for all sub-samples as well as the full sample indicating a significant positive relationship between metacognition and eudaimonic well-being.

Hypothesis 4

Hypothesis 4 states that goal-striving does not significantly predict eudaimonic well-being over and above metacognition. In the full sample, the GLM approach indicated that goal-striving still significantly explained 1.8% of variance in eudaimonic well-being, $F(1, 508) = 14.59$, $B = .43$, $\beta = .15$, $p < .001$, when controlling for the effect of metacognition, $F(1, 509) = 219.63$, $\Delta R^2 = .281$, $R^2 = .349$, adjusted $R^2 = .345$, $B = .107$, $\beta = .48$, $p < .001$. The regression coefficient also indicated that for every 1 SD change in GS, EWB changes by .15 SD. The results for the sub-samples are presented in Table 5.

Although the unique contribution of goal-striving to EWB substantially dropped after controlling for metacognition (the mediator) indicating partial mediation, a full mediation was not supported for the full sample and for any of sub-samples except for adults, $F(1, 81) = .555$, $\Delta R^2 = .004$, $B = .17$, $\beta = .07$, $p = .46$. The

Table 4. means and standard deviations for eudaimonic well-being (EWB), metacognition (MC), goal-striving (GS), and goal-aspiration (GA) for the full sample and for sub-samples

| (sub-)samples | variables of interest | | | | | | | |
| | EWB | | MC | | GS | | GA | |
	M	SD	M	SD	M	SD	M	SD
male ($n = 183$)	59.84	12.46	337.31	53.48	32.61	4.59	76.34	9.46
female ($n = 330$)	62.46	12.04	340.15	56.63	34.23	3.85	79.48	8.64
significant mean differences	no		no		yes		yes	
education ($n = 312$)	63.56	12.06	342.78	55.59	34.41	3.89	79.57	9.17
non-education ($n = 201$)	58.36	11.88	333.48	55.00	32.48	4.39	76.47	8.57
significant mean differences	yes		no		yes		yes	
adult ($n = 85$)	66.92	10.98	356.46	51.51	33.76	4.35	80.10	9.48
emerging adults ($n = 428$)	60.45	12.21	335.70	55.67	33.63	4.17	78.01	8.94
significant mean differences	yes		yes		no		no	
full sample ($N = 513$)	61.52	12.24	339.14	55.49	33.65	4.20	78.36	9.06

reduction in ΔR^2 was the greatest for males (.118) and followed closely by non-education students (.113) and, was the smallest for adults (.058), though there was a full mediation. That is that in adults the whole variance explained by GS in EWB was accounted for by their metacognition. Overall, Hypothesis 4 was only supported for the subsample of adults.

Hypothesis 5

Hypothesis 5 states that the indirect effect of goal-striving on eudaimonic well-being (i.e., the contribution of goal-striving to eudaimonic well-being through metacognition) is significant. The indirect effect of the independent variable on the dependent variable through the mediator, in this case the indirect effect of Goal-Striving on EWB through Metacognition, is known as the index of mediation. AMOS 18.0 was used to test the mediation model and to indicate the significance of the indirect effect. For the full sample and for each subsample, important demographic variables (i.e., age group, major, and gender) were entered into the models as appropriate. All models were investigated for model fit using global and focused fit indices (i.e., chi-square, the Root Mean Square Error of Approximation, the p value for the close fit, the Comparative Fit Index, the Standardized Root Mean Residual, the standardized residual covariance, and modification indices), and fit indices indicated good fit for all models. Using AMOS 18.0 for Shrout and Bolger's (2002) Bootstrap procedure, the researchers requested 5000 bootstrap samples with 95% confidence interval (bias-corrected percentile method). The result for the mediation index is presented in Table 6.

The results indicated that the indirect effect of Goal-Striving on EWB (i.e., the mediational effect) was statistically significant for the full sample and for all sub-samples. For instance, the results for the full sample was $B = .5094$ (95% CI [.3799, .6672]), $\beta = .1748$ (95% CI [.1317, .2236]), $p < 0.001$. That is, on average for every 1 SD positive change in Goal-Striving, EWB changes by .17 SD. This model explained a total variance of 36% in EWB. For Adults, $B = .4965$ (95% CI [.1761, .9228]), $\beta = .2007$ (95% CI [0.0754, 0.3572]), $p < 0.001$, and for emerging adults, $B = .5078$ (95% CI [.3721, .6721]), $\beta = .1730$ (95% CI [.1291, .2261]), $p < 0.001$,

Table 5. summary of general linear model predicting eudaimonic well-being from goal-striving while controlling for metacognition

variable	eudaimonic well-being				
full sample ($N = 513$)	β^*	ΔR^2	ΔF	df_1	df_2
age group and major (block 1)	—	0.068	18.538	2	510
metacognition	0.483¶	—	—	—	—
block 2	—	0.281	219.63¶	1	509
goal-striving	0.149¶	—	—	—	—
block 3	—	0.018	14.593¶	1	508
total $R^2 = 0.367$, $\bar{R}^2 = 0.362$					
female ($n = 330$)	β^*	ΔR^2	ΔF	df_1	df_2
age group and major (block 1)	—	0.071	12.586¶	2	327
metacognition	0.490¶	—	—	—	—
block 2	—	0.273	135.78¶	1	326
goal-striving	0.122†	—	—	—	—
block 3	—	0.013	6.560†	1	325
total $R^2 = 0.357$, $\bar{R}^2 = 0.350$					
male ($n = 183$)	β^*	ΔR^2	ΔF	df_1	df_2
age group and major (block 1)	—	0.044	4.157†	2	180
metacognition	0.491¶	—	—	—	—
block 2	—	0.308	85.22¶	1	179
goal-striving	0.171†	—	—	—	—
block 3	—	0.023	6.567†	1	178
total $R^2 = 0.375$, $\bar{R}^2 = 0.361$					
non-education students ($n = 201$)	β^*	ΔR^2	ΔF	df_1	df_2
age group (block 1)	—	0.000	0.015	1	199
metacognition	0.491¶	—	—	—	—
block 2	—	0.310	89.14¶	1	198
goal-striving	0.157†	—	—	—	—
block 3	—	0.020	5.965†	1	197
total $R^2 = 0.331$, $\bar{R}^2 = 0.321$					
education students ($n = 312$)	β^*	ΔR^2	ΔF	df_1	df_2
age group (block 1)	—	0.050	16.40¶	1	310
metacognition	0.488¶	—	—	—	—
block 2	—	0.276	126.38¶	1	309
goal-striving	0.144‡	—	—	—	—
block 3	—	0.019	8.694‡	1	308
total $R^2 = 0.344$, $\bar{R}^2 = 0.338$					
adults ($n = 85$)	β^*	ΔR^2	ΔF	df_1	df_2
major (block 1)	—	0.111	10.343‡	1	83
metacognition	0.541¶	—	—	—	—
block 2	—	0.294	40.510¶	1	82
goal-striving	0.068	—	—	—	—
block 3	—	0.004	0.555	1	81
total $R^2 = 0.409$, $\bar{R}^2 = 0.387$					
emerging adults ($n = 428$)	β^*	ΔR^2	ΔF	df_1	df_2
major (block 1)	—	0.023	10.236‡	1	426
metacognition	0.474¶	—	—	—	—
block 2	—	0.287	176.55¶	1	425
goal-striving	0.170¶	—	—	—	—
block 3	—	0.023	14.884¶	1	424
total $R^2 = 0.333$, $\bar{R}^2 = 0.329$					

* The standardized regression coefficients are reported from the overall model and only for the main variables. † $p < 0.05$; ‡ $p < 0.01$; ¶ $p < 0.001$.

indicating a significant mediational effects. That is, on average for every 1 SD positive change in Goal-Striving, EWB changes by .20 SD for adults and by .17 SD for emerging adults. The mediational model explained a total variance of 38% in EWB of adults and 34% in EWB of emerging adults. Overall, the result of bootstrap procedure supported the mediational effect of metacognition in the relationship between goal-striving and eudaimonic well-being. This mediational effect was partial for all sub-samples except for sub-samples of adults. For adults, the mediational model showed a greater effect size and indicated a full mediation.

Mediated variance
To illustrate an approximation of the unique mediational contribution of goal-striving (GS) to eudaimonic well-being (EWB) through metacognition (MC), the estimates of R-Squared Change for the unique contribution of GS to EWB before and after controlling for MC was subtracted from each other ($\Delta R^2_{without\ mediator} - \Delta R^2_{with\ mediator}$). The result, presented in Table 7, indicates the proportion of variance in EWB accounted for by mediation or, in other words, the mediated variance shared by the three variables. The results indicated that the mediated variance was 8.3% for the full sample. The greatest mediated variance was detected for males (11.8%) closely followed by non-education students (11.3%), and it was the smallest for adults (5.8%).

Hypothesis 6
Hypothesis 6 states that goal-aspiration (GA) moderates the relationship between goal-striving (GS) and eudaimonic well-being (EWB). To produce the interaction term, scores on goal-striving and goal-aspiration were standardized to avoid possible multicollinearity problems (Stevens, 2002). In the full sample, the GLM approach indicated that the interaction between GS and GA, which reflects the moderation effect, accounted for a small yet statistically significant percentage of variance in EWB, $F(1, 507) = 4.120$, $\Delta R^2 = .006$, $R^2 = .219$, $adjusted\ R^2 = .212$, $B = -.87$, $\beta = -.08$, $p < .05$. Similar results, with a slightly larger effect size, were observed for adults, $F(1, 80) = 4.53$, $\Delta R^2 = .044$, $R^2 = .218$, $adjusted\ R^2 = .179$, $B = -1.93$, $\beta = -.21$, $p < .05$, but not for emerging adults or any of the other subsamples. The results are presented in Table 8. Therefore, Hypothesis 6 was supported for the full sample and for the sub-sample of adults only.

Moderating effect of goal-aspiration
To plot the interaction, the regression analysis was rerun using the dichotomous GA and standardized scores for GS to obtain the regression coefficients. A median split was used to dichotomize the goal-aspiration variable. Scores at or above the median, 78.00 were categorized as the high goal-aspiration group and scores below 78.00 labeled as the low goal-aspiration group.

Table 6. summary of mediation analyses predicting eudaimonic well-being from goal-striving (GS) with metacognition (MC) as the mediator

(sub-)samples	indirect effect of GS on EWB		
	point of estimate	lower−upper bound (95% CI)	overall R^2 (lower−upper bound)
full sample ($N = 513$)	0.5094 (0.1748)	0.3799−0.6672 (0.1317−0.2236)†	0.3657* (0.2874−0.4281)
female ($n = 330$)	0.4854 (0.1550)	0.3134−0.6989 (0.1013−0.2169)†	0.3598* (0.2644−0.4426)
male ($n = 183$)	0.5948 (0.2199)	0.4061−0.8300 (0.1517−0.3026)†	0.3707* (0.2553−0.4656)
non-education ($n = 201$)	0.5611 (0.2074)	0.3791−0.7963 (0.1414−0.2868)†	0.3310* (0.2123−0.4313)
education ($n = 312$)	0.4785 (0.1541)	0.3032−0.6967 (0.0993−0.2182)†	0.3479† (0.2517−0.4308)
adult ($n = 85$)	0.4965 (0.2007)	0.1761−0.9228 (0.0754−0.3572)†	0.3842* (0.1650−0.5344)
emerging adults ($n = 428$)	0.5078 (0.1730)	0.3721−0.6721 (0.1291−0.2261)†	0.3374* (0.2571−0.4089)

Note. Significance levels are based on p-values estimated by bias-corrected confidence intervals. Standardized regression coefficients are in parentheses. * $p < .01$; † $p < .001$.

Plotting the interaction effect indicated that for the full sample (see Figure 1), while the relationship between goal-striving and eudaimonic well-being is positive and statistically significant for both high and low goal-aspiration groups , higher goal-aspiration has a positively moderating effect on this relationship. That is, students with higher level of eudaimonic self-actualizing aspiration reported higher goal-striving and higher eudaimonic well-being compared to students with lower level of eudaimonic self-actualizing aspirations.

On the other hand, for the adult sub-sample, the relationship between GS and EWB was not significant for the high level of goal-aspiration, but it was positive and significant for students with the lower level of goal-aspiration. As illustrated in Figure 2, while the maximum level of eudaimonic well-being for adult students with high eudaimonic self-actualizing aspirations was steady around the observed group mean (68.67) and did not exceed 70, eudaimonic well-being of adult students with low goal-aspiration exceeded that average point as the scores on the quality of their goal striving increased. That is, while low goal-aspiration group were lower in their eudaimonic well-being at the lower levels of goal-striving compared to high goal-aspiration, as their goal-striving exceeded the sample mean (33.65) their eudaimonic well-being also exceeded eudaimonic well-being of high goal-aspiration group.

Goal-aspiration, goal-striving and eudaimonic well-being in remaining sub-samples

Conducting GLM, where GA was entered in the first step and GS in the second step, indicated that goal-aspiration positively and significantly predicted EWB. Similarly, goal-striving also positively and significantly predicts eudaimonic well-being for all remaining sub-samples over and above goal-aspiration. For non-education students, goal-aspiration significantly accounted for 13.0% of variance in eudaimonic well-being, and goal-striving accounted for additional 5.9% of unique variance; for education students the amount of variance accounted for by GA and GS respectively

was 9.6% and 3.1%. For female students 6.8% and 3.4%, for male students 17.7% and 5.1%, and for emerging adults 12.6% and 4.6%. Table 9 present the results of these analyses.

Overall, Hypothesis 6 was supported for the full sample and for the sub-sample of adult students, meaning, goal-aspiration moderated the relationship between goal-striving and eudaimonic well-being. In the full sample, students with high eudaimonic self-actualizing aspiration reported higher level of well-being at all reported levels of goal-striving compared to students with low eudaimonic self-actualizing aspiration. For adults, EWB was steady for the high goal-aspiration group while it was positively increasing for the low goal-aspiration group by increases in goal-striving. In other sub-samples, there was no interaction effect, but goal-striving accounted for unique variances in EWB over and above goal-aspiration. The greatest variance (in EWB) accounted for by GA was for males (17.7%) and the smallest variance was observed for females (6.8%). The amount of unique variance (in EWB) shared by GS over and above that of GA did not exceed 6.0% for any of the sub-samples. This amount of variance is notably lower than what was observed in testing Hypothesis 1 where there was no consideration of goal-aspiration (also see Table 7 for ΔR^2 $_{without}$ $_{mediator}$). This indicated that a percentage of the variance in EWB is shared by both goal-aspiration and goal-striving. Given the large amount of shared variance by GA and GS, it could be expected that metacognition fully mediates the relationship between GS and EWB when GA is being used as the moderator. This prediction was partially supported by the results of the moderated mediation analyses for testing Hypothesis 7.

Hypothesis 7

Hypothesis 7 states that goal-aspiration (GA) moderates the mediated relationship between goal-striving (GS) on eudaimonic well-being (EWB) where metacognition is the mediator. To test Hypothesis 7 for the full sample and the sample of adults, which already (in Hypothesis 6) has shown a significant interaction effect between GA and GS, a multigroup approach was used with the same SEM model as in Hypothesis 5. For the remaining sub-samples, goal-aspiration was added to their respective SEM models. Testing the models for fit indices indicated a good fit (Chi-square = 6.8, $df = 7$, $p > .05$; CFI $\geq .95$, $P_{close} \geq .05$; RMSEA $\leq .08$) for all models. Then, mediation was tested for the full sample and for the sub-samples (Table 9).

Table 7. the mediated variance shared by goal-striving, metacognition, and eudaimonic well-being

(sub-)samples	unique direct contribution of GS to EWB		
	ΔR^2 without mediator	ΔR^2 with mediator	% of mediated variance
full sample ($N = 513$)	0.101	0.018	8.3
female ($n = 330$)	0.078	0.013	6.5
male ($n = 183$)	0.141	0.023	11.8
non-education ($n = 201$)	0.133	0.020	11.3
education ($n = 312$)	0.089	0.019	7.0
adult ($n = 85$)	0.058	0.000	5.8
emerging adults ($n = 428$)	0.116	0.023	9.3

Table 8. summary of GLM predicting EWB from goal-striving and goal-aspiration

variable	eudaimonic well-being				
full sample (N = 513)	**β***	**ΔR²**	**ΔF**	**df₁**	**df₂**
age group and major (block 1)	—	0.068	18.54¶	2	510
goal-aspiration	0.238¶	—	—	—	—
goal-striving	0.213¶	—	—	—	—
block 2	—	0.145	46.82¶	2	508
GA × GS	−0.081†				
block 3	—	0.006	4.120†	1	507
total R² = 0.219, R̄² = 0.212					
adults (n = 85)	**β***	**ΔR²**	**ΔF**	**df₁**	**df₂**
major (block 1)	—	0.111	10.34‡	1	83
goal-aspiration	0.074	—	—	—	—
goal-striving	0.211	—	—	—	—
block 2	—	0.063	3.103†	2	81
GA × GS	−0.212†	—	—	—	—
block 3	—	0.044	4.527†	1	80
total R² = 0.218, R̄² = 0.179					
emerging adults (n = 428)	**β***	**ΔR²**	**ΔF**	**df₁**	**df₂**
major (block 1)	—	0.023	10.24‡	1	426
goal-aspiration	0.026	—	—	—	—
block 2	—	0.126	63.14¶	1	425
goal-striving	0.241	—	—	—	—
block 3	—	0.046	24.16¶	1	424
total R² = 0.196, R̄² = 0.190					
non-education students (n = 201)	**β***	**ΔR²**	**ΔF**	**df₁**	**df₂**
age group (block 1)	—	0.000	0.015	1	199
goal-aspiration	0.258¶	—	—	—	—
block 2	—	0.130	29.52¶	1	198
goal-striving	0.264¶	—	—	—	—
block 3	—	0.059	14.43¶	1	197
total R² = 0.189, R̄² = 0.177					
education students (n = 312)	**β***	**ΔR²**	**ΔF**	**df₁**	**df₂**
age group (block 1)	—	0.050	16.40¶	1	310
goal-aspiration	0.221¶	—	—	—	—
block 2	—	0.096	34.75¶	1	309
goal-striving	0.199‡	—	—	—	—
block 3	—	0.031	11.77‡	1	308
total R² = 0.178, R̄² = 0.170					
female (n = 330)	**β***	**ΔR²**	**ΔF**	**df₁**	**df₂**
age group and major (block 1)	—	0.071	12.59¶	1	327
metacognition	0.176¶	—	—	—	—
block 2	—	0.068	25.80¶	1	326
goal-striving	0.207¶	—	—	—	—
block 3	—	0.034	13.42¶	1	325
total R² = 0.174, R̄² = 0.164					
male (n = 183)	**β***	**ΔR²**	**ΔF**	**df₁**	**df₂**
age group and major (block 1)	—	0.044	4.157†	1	180
metacognition	0.323¶	—	—	—	—
block 2	—	0.177	40.59¶	1	179
goal-striving	0.251‡	—	—	—	—
block 3	—	0.051	12.42‡	1	178
total R² = 0.272, R̄² = 0.255					

* The standardized regression coefficients are reported from the overall model and only for the main variables. † $p < 0.05$; ‡ $p < 0.01$; ¶ $p < 0.001$.

The result of the mediation analyses indicated the full mediation for high goal-aspiration group in the full sample (Direct effect = .2227, p = .18; Indirect effect = .3129, $p < .01$ while for low-goal-aspiration group in the full sample only partial mediation was detected (Direct effect = .4670, $p < .01$; Indirect effect = .5171, $p < .001$). In addition, full mediation was supported for the low goal-aspiration group in the adult sub-sample (Direct effect = .4938, p = .38; Indirect effect = .7790, p < .05). However, no mediational effect was detected in adults with high goal-aspiration nor was there a direct relationship between goal-striving and EWB, indicating no relationships between goal-striving and EWB.

As expected, partial mediational effects for most of sub-samples changed to full mediational effect when goal-aspiration was added to the model. Altogether the results for all subsamples, and in general for the students with high goal-aspirations, were in line in supporting the hypothesis 7 that stated goal-aspiration moderates the mediational effect of metacognition in the relationship between GS and EWB. Most of the partial mediations in the subs-samples became full mediations except for emerging adults and for adults with high goal-aspiration which no direct or indirect effect was observed.

Summary of the results

In sum, Hypotheses 1 to 3 were supported for the full sample and for all the subsamples. The individuals who scored higher on goal-striving were more likely to score higher on eudaimonic well-being. This suggested that individuals with higher quality of goal-striving experience higher level of eudaimonic well-being. In addition, the individuals who scored higher on goal-striving were more likely to score higher on general metacognitive competence. This implied that individuals with higher quality of goal-striving also demonstrate higher general metacognitive competence. Further, the individuals who scored higher on metacognition were more likely to score higher on eudaimonic well-being, suggesting that individuals with higher general metacognitive competence experience higher level of eudaimonic well-being.

On the other hand, Hypothesis 4 was only supported for the sub-sample of adults. Comparing adults and emerging adults, the finding suggests that for adults the relationship of goal-striving with eudaimonic well-being is fully dependent on their metacognition. However, for emerging adults the quality of goal-striving contributes to their eudaimonic well-being above and beyond their general metacognitive competence. The analyses on the other sub-samples also yielded similar results as that of emerging adults. The follow-up mediation analyses in Hypothesis 5 also supported the full mediation for adults but not for any other sub-samples. While this mediational relationship was more dominant for adults than it was for emerging adults (i.e., the full mediation vs. partial mediation), metacognition and goal-striving shared a greater amount of variance in emerging adults (9.3%) compared to adults (5.8%). This
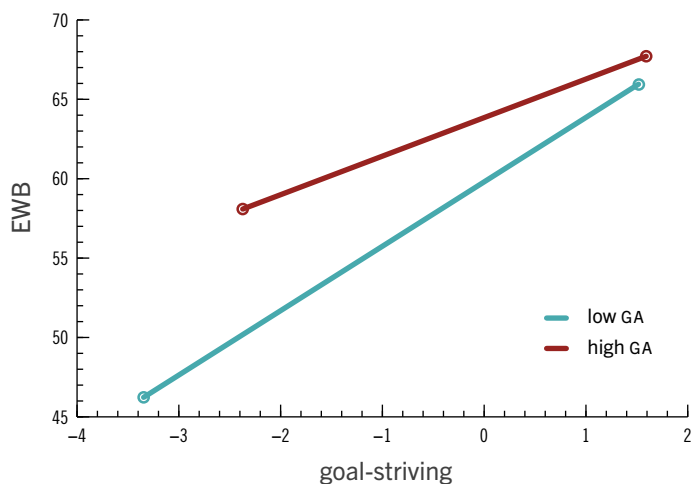
Figure 1. the goal-aspiration by goal-striving interaction in predicting eudaimonic well-being for the full sample

finding confirms that for both cases, higher quality of goal-striving is linked to higher metacognition which fosters higher eudaimonic well-being. It also indicates that goal-striving of emerging adults, while sharing metacognition as the common element, is different from adults' in defining eudaimonic well-being.

Hypotheses 6 and 7 took goal-aspiration into account as a moderator. While an interaction was detected for the full sample and for the sub-sample of adults, there was no interaction for the remaining sub-samples. For the full sample, goal-aspiration moderated the relationship between goal-striving and eudaimonic well-being in a way that individuals with the higher level of goal-aspiration experienced higher level of eudaimonic well-being at all reported levels of goal-striving compared to individuals with lower level of goal-aspiration. Adults with high aspiration remained steady in their eudaimonic well-being regardless of their goal-striving. However, goal-striving positively predicted eudaimonic well-being for adults with low goal-aspiration indicating that the higher the level of goal-striving is for these individuals, the higher their eudaimonic well-being would become regardless of their low goal-aspiration.

In addition, the observed full mediations after goal-aspiration was taken into consideration indicated that goal-aspiration and metacognition together explained all the shared variance between EWB and goal-striving (except for low-aspiration students and emerging adults). This finding partly suggests that goal-aspiration completes the mediational effect of metacognition. Altogether, metacognition and goal-aspiration were found to be two overarching constructs in contributing to individuals' eudaimonic well-being.

» DISCUSSION

Virtues of character and intellect, as discussed by Aristotle (2004), are two main elements responsible for individuals' well-being. Achieving excellences in these two kinds of virtue allows individuals to perform at the optimal level of rational choice and exercise of reason when making decisions in their everyday lives, in setting their goals and in their efforts in pursuing their goals. This optimal functioning creates a life of *acting well and living well*. We pre-

sented a rationale that contemporary concepts of goal-aspiration and metacognition can represent these two kinds of virtue, and different levels of goal-aspiration and metacognition can indicate different levels of character and intellect. Further, we argued that different levels of this optimal functioning can be represented within a contemporary conceptualization of self-actualization in which self-actualization is a goal-oriented process (Rule, 1991) and can in part be represented by two concepts of goal-aspiration and goal-striving. Different levels of self-actualization, then, argued to be associated with different levels of eudaimonic well-being, a well-being conception which represents the ultimate good in a subjective way.

Although none of the previous studies of well-being, self-actualization, or metacognition has employed the conceptual framework presented in this study, there were some research studies that investigated the links between different representations of these constructs. For instance, the positive link between self-actualization and subjective well-being and psychological well-being has been detected in the literature where self-actualization was argued to be represented by the highest level of Loevinger's (1976) ego development (Bauer, Schwab, & McAdams, 2011). Previous research on goal-related variables and individuals' well-being also indicated the positive effect of intrinsic and autonomous (versus extrinsic and controlled) goal aspirations and goal attainment, and goal progress on self-actualization (measured as a personality characteristic) and subjective and psychological well-being (e.g., Kasser & Ryan, 1996, 2001; Sheldon & Kasser, 1998; Sheldon, Kasser, Smith, & Share, 2002; Sheldon, Ryan, Deci, & Kasser, 2004). Further, research indicated a link between self-awareness (an aspect of metacognition) and self-actualization and self-actualizing personalities (e.g., Bar-on, 2001, 2006; Culbert, Clark, & Bobele, 1968). Studies, also, indicated a link between cognitive and metacognitive efficiency and psychological well-being (e.g., Fastame, Penna, Rossetti, & Agus, 2013).

Following our argument, we developed an empirical study that examined these hypothesized relationships with a focus on exploring the mediational role that metacognition (as a way of exercise of goal-oriented reason and as a teachable skill) may play
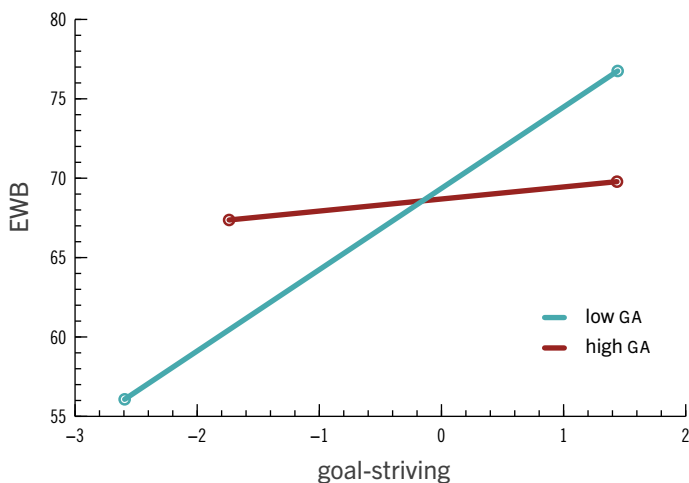

Figure 2. the goal-aspiration by goal-striving interaction in predicting eudaimonic well-being for the subsample of adults

Table 9. summary of moderated mediation analyses predicting eudaimonic well-being from goal-striving (GS) with metacognition (MC) as the mediator and goal-aspiration (GA) as the moderator

| (sub-)samples | unique direct contribution of GS to EWB | | |
| | without mediator | with mediator | |
	direct effect	direct effect	indirect effect
full sample (*N* = 513)			
high GA (n = 201)	sig	0.2227 (0.0723)	0.3129 (0.1015)†
low GA (n = 312)	sig	0.4670 (0.1621)†	0.5171 (0.1795)‡
adult (*n* = 85)			
high GA (n = 50)	ns	−0.0278 (−0.0108)	0.2419 (0.0937)
low GA (n = 35)	sig	0.4938 (0.1900)	0.7790 (0.2998)*
emerging adults (*n* = 428)	sig	1.0731 (0.0878)*	2.0247 (0.1656)‡
female (*n* = 330)	sig	0.9176 (0.0707)	1.9955 (0.1525)‡
male (*n* = 183)	sig	1.0575 (0.0927)	2.3202 (0.2034)‡
education (*n* = 312)	sig	0.9096 (0.0704)	1.8371 (0.1421)‡
non-education (*n* = 201)	sig	1.0997 (0.0968)	2.1967 (0.1934)‡

Note. Only unstandardized regression coefficients were reported for direct and indirect effects. * $p < 0.05$; † $p < 0.01$; ‡ $p < 0.001$.

in the relationship between self-actualizing efforts and eudaimonic well-being. Findings from the present study suggest that the quality goal-striving, which was indicated by higher goal-importance and commitment, and higher effort, initiation and inspiration for pursuing personal goals, make a positive contribution to individuals' well-being. The findings also suggest that individuals' general metacognitive competence, as operationally defined and measured in this study, was largely responsible for this contribution. While this effect was more dominant for adults than it was for emerging adults, the finding implies that, for both cases, the development of general metacognitive competence provides the internal competency for a quality goal pursuit that, in turn, helps with higher experience of eudaimonic well-being.

In other words, metacognition enhances eudaimonic well-being at least partly through its practical effect manifested in the quality of goal pursuit. The metacognitive competence (including monitoring, evaluation, goal-setting, planning, strategy use, and strategy adjustment) provides the intrapersonal skills required for a quality goal-striving that enhances eudaimonic well-being. That is, metacognitive thinking provides a medium to exercise high-quality goal-pursuits and self-actualization that nurture eudaimonic well-being. Nevertheless, the contribution of metacognition to eudaimonic well-being was beyond its association with goal-striving.

In addition, goal-aspiration as an indicator of the eudaimonic aspects of self-actualization showed to be influential in students' well-being as they strive for pursuing their personally worthwhile goals. The way individuals strive for their goals and the reason they pursue those goals has some shared effects on their eudaimonic well-being. But the findings from this study imply that the exercise of reason and character seems to play an overarching role when it comes to individuals' goal pursuit and well-being. That is, the role of eudaimonic self-actualizing aspiration was complementary to general metacognitive competence in contributing to those aspects of eudaimonic well-being associated with goal-striving.

Yet, in general the relationship between goal-striving and eudaimonic well-being is more complex for students with lower eudaimonic self-actualizing aspirations – thus, with less-than-optimal rational choices – compared to those with higher goal-aspiration. We can argue that in addition to the intellectual aspects of their goal-pursuit, either there may be other aspects to the goal-striving of these groups that contribute to their well-being, or their eudaimonic well-being enhances by a sense of meaning and engagement they may directly develop through committing to personally inspiring effortful planned goal pursuits. Similar complexity exists for emerging adults, who regardless of their goal-aspiration and the intellectual aspects of their goal-striving find other eudaimonic fulfillments in their pursuits.

Implications
Metacognition is an information-processing competency that is manifested more clearly in problem-solving tasks and situations, and is known as an attainable competence; that is, it can be taught as sets of skills for goal-setting and problem-solving (Borkowski & Muthukrishna, 1992; Gaskins & Pressley, 2007). The finding of this study implies that developing general metacognitive skills prepares individuals for a quality goal-striving and helps them experience higher levels of eudaimonic well-being in part by an effective and enhanced goal-oriented effort. Thus, teaching metacognition as early as possible may equip our students and individuals with an essential competence for paving their ways towards their optimal functioning and eudaimonic well-being.

Further, goal-aspiration is more or less a matter of character and a trained desire to appreciate eudaimonic ends, namely, the ends which are intrinsic, constitutive, self-concordant, and personally expressive and particularly value-oriented. As Aristotle (2004) argued, virtue of character is mainly developed through habituation and training from early ages. He argued that by being surrounded by right rules and laws, children may be able to train their desires in a right direction as they practice the rational choice under these rules and laws. Character education as well as right parenting may play influential roles here. However, as Aristotle argued, to build a culture for good character, developing a *good* legislation system may be necessary.

Because "rational choice is either desire-related intellect or thought-related desire" (p.105), it still requires the development of both intellect and character to function at its optimal level. Thus, this two virtues need to be addressed in parallel as one lacking the other may not function as good. An intellectual person who lacks character and morality could turn into a cruel clever person. The virtuous person lacking intellect may not be able to adequately differentiate between right and wrong in varying situations and, thus, be unable to make or perform a practically wise choice.

Directions for future research

The finding implies that goal-strivings of adults and emerging adults are different in nature and quality as their pattern of contribution to eudaimonic well-being is different. While metacognition is responsible for adults' actualizing efforts, the quality of goal-striving that contributes to emerging adults' eudaimonic well-being is only partly related to their general metacognitive competence. This suggests that there may be some other intrapersonal, interpersonal, and environmental factors in the life of emerging adults that explain this discrepancy. The researchers recommend further exploration of the nature of goal pursuits for emerging adults in comparison with adults to add insight into this discord.

Furthermore, the finding suggests that developing metacognition and pursuing eudaimonic self-actualizing aspirations both have advantages in creating a eudaimonic living that go beyond a quality goal-striving. The metacognitive competence may help with the goal selection (i.e., the identification of the problem or purpose and the kind of goals one selects to pursue that purpose or to solve the problem), which is a step prior to goal-striving. Being metacognitive also means acting with more self-awareness (Brown, 1977; Brown & Smiley, 1977), which may be the link to the personal expressiveness and eudaimonic well-being. Further, high goal-aspiration indeed represents some character strengths that may be in part responsible for the enhancement of eudaimonic well-being. Scrutinizing the complex nature of the relationship between eudaimonic well-being and these two constructs and discovering ways in which developing metacognition and character can contribute to eudaimonic well-being is a subject for future research.

In sum, metacognition and goal-aspirations are only two elements that were found to contribute to eudaimonic well-being as they are related to individuals' goal-oriented pursuits. There is a great amount of variance in eudaimonic well-being, which this study was unable to explain. Given the great concern people have over living a fulfilling life, research into other possible contributors to individuals' well-being will be enlightening. Self-actualization and metacognition are also complex constructs, research into which can be of merit for their theoretical and practical implications in the field of education, psychology, and in life. ■

REFERENCES

Andrews, F. M., & Withey, S. B. (1976). *Social indicators of well-being.* New York, NY: Plenum Press.

Aristotle (2004). *Nicomachean ethics* (R. Crisp, Trans.). New York, NY: Cambrige University Press. (Original work in 4th Century BC).

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist, 55*, 469-480.

Bar-On, R. (2001). Emotional intelligence and self-actualization. In J. Ciarrochi, J. Forgas, and J. D. Mayer (Eds.), *Emotional intelligence in everyday life: A scientific inquiry*. New York, NY: Psychology Press.

Bar-On, R. (2006). The Bar-On model of emotional-social intelligence (ESI). *Psicothema*, *18*, supl., 13-25.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Bauer, J. J., Schwab, J. R., & McAdams, D. P. (2011). Self-actualizing: Where ego development finally feels good? *The Humanistic Psychologist*, *39(2)*, 121-136.

Borkowski, J. & Muthukrishna, N. (1992). Moving metacognition into the classroom: "Working models" and effective strategy teaching. In Pressley, Michael (Ed); Harris, Karen R. (Ed); Guthrie, John T. (Ed), *Promoting academic competence and literacy in school* (pp. 477-501). San Diego, CA, US: Academic Press.

Brown, A. L. (1977). Development, schooling, and the acquisition of knowledge about knowledge. In R. C. Anderson, R. J. Spiro, & W. E. Montague (be Eds.), *Schooling and the acquisition of knowledge* (pp. 241-253). Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, A. L. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. E. Weinert, & R. H. Kluwe (Eds.), *Metacognition, motivation and understanding* (pp. 65-116). Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, A. L., & Smiley, S. S. (1977). Rating the importance of structural units of prose strategies: A problem of metacognitive development. *Child Development, 48*, 1-8.

Culbert, S. A., Clark, J. V., & Bobele, H. K. (1968). Measures of change toward self-actualization in two sensitivity training groups. *Journal of Counseling Psychology, 16*(1)*, 53-57.

Debats, D. L. (1998). Measurement of personal Meaning: the psychometric Properties of the Life Regard Index. In T. P. Wong and P. S. Fry (Eds.), *The human quest for meaning: A hand book of psychological research and clinical applications* (pp.237-260). Mahwah, NJ: Lawrence Erlbaum Associates.

Deci, E. L., & Ryan, R. M. (2008). Hedonia, eudaimonia, and well-being: an Introduction. *Journal of Happiness Studies, 9*, 1-11.

Diener, E. (1984). Subjective well-being. *Psychological Bulletin, 95*, 542-575.

Diener, E. Emmons, R. A., Larsen, R. J. and Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*, 71-75.

Diener, E., & Lucas, R. E. (1999). Personality and subjective well-being. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 213-229). New York, NY: Russell-Sage.

Diener, E., Suh, M., Lucas, E. & Smith, H. (1999). Subjective well-being: Three decades of progress, *Psychological Bulletin, 125*(2), 276-302.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York, NY: Chapman & Hall.

Emmons, R. A. (1999). *The psychology of ultimate concerns*. Motivation and spirituality in personality. New York: Guilford.

Fastame, M. C., Penna, M. P., Rossetti, E. S., & Agus, M. (2013). Perceived well-being and metacognitive efficiency in life course: A developmental perspective. *Research on Aging, 35* (6). 736-749.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist, 34,* 906 - 911.

Fowers, B. J., Mollica, C. O., & Procacci, E. N. (2010).Constitutive and instrumental goal orientations and their relations with eudaimonic and hedonic well-being. *The Journal of Positive Psychology, 5*(2), 139 153.

Gaskins, I. W. & Pressley, M. (2007). Teaching metacognitive strategies that address executive function processes within a school-wide curriculum. In L. Meltzer (Ed.) *Executive function in education: From theory to practice* (pp. 261-286). New York, NY: The Guildford Press.

Jones, R., & Crandall, R. (1986). Validation of a short index of selfactualization. *Personality and Social Psychological Bulletin, 12,* 6 3-73.

Kasser, T., & Ryan, R. M. (1996). Further examining the American dream: Differential correlates of intrinsic and extrinsic goals. *Personality and Social Psychology Bulletin*, *22,* 280-287.

Kasser, T., & Ryan, R. M. (2001). Be careful what you wish for: Optimal functioning and the relative attainment of intrinsic and extrinsic goals. In P. Schmuck and K. M. Sheldon (Eds), *Life Goals and Well-Being: Towards a Positive Psychology of Human Striving* (pp. 115–129). Goettingen, Germany: Hogrefe and Huber Publishers.

Kerr, B. A. (1991). Educating gifted girls. In Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 402-415). Neeham Heights, MN: Allyn & Bacon.

Kerr, B. A. (1985). *Smart girls, gifted women.* Dayton, Ohio: Ohio Psychology Press.

Kraut, R. (1979). Two conceptions of happiness. *Philosophical Review, 87*, 167–196.

Lin, L. M. & Zabrucky, K. M. (1998). Calibration of comprehension: Research and implications for education and instruction. *Contemporary Educational Psychology, 23*, 345-391.

Loevinger, J. (1976). *Ego development*. San Francisco, CA: Jossey-Bass.

MacKinnon, D. P., Lockwood, C., Hoffman, J. West, S. & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*, 83-104.

Maslow, A. H. (1968). *Towards a psychology of being* (2nd ed.). New York, NY: Van Nostrand Reinhold.

McCormick, C. B. (2003). Metacognition and learning. In W. M. Reynolds & G. E. Miller (Eds*.*), *Handbook of psychology: Educational psychology* (pp. 79-102). Hoboken, NJ: John Wiley & Sons.

Milyavskaya, M., Ianakieva, I., Foxen-Craft, E., Colantuoni, A., & Koestner, R. (2012). Inspired to Get There: The Effects of Trait and Goal Inspiration on Goal Progress. *Personality and Individual Differences, 52*, 56-60.

Mooney, C., & Duval, R. (1993). *Bootstrapping.* Newbury Park, CA: Sage.

Ng, K. Y., Ang, S., Chan, K. Y. (2008). Personality and leader effectiveness: A moderated mediation model of leadership self-efficacy, job demands, and job autonomy. *Journal of Applied Psychology, 93*(4), 733-43.

Norton, D. L. (1976). *Personal destinies.* Princeton, NJ: Princeton University Press.

Peterson, C., Park, N., & Seligman, M. E. (2005). Orientations to happiness and life satisfaction: The full life versus the empty life. *Journal of Happiness Studies, 6*, 25-41.

Reis, S. M., & Callahan, C. (1989). Gifted females: They've come a long way--or have they. *Journal for the Education of the Gifted, 12* (2), 99-117.

Rindfuss, R. R. (1991). The young adult years: Diversity, structural change, and fertility. *Demography, 28*, 493-512.

Robitschek, C. (1998). Personal Growth Initiative: The construct and its measures. *Measurement and Evaluation in Counseling and Development, 30*(4), 183-198.

Rogers, C. R. (1951). *Client-Centered therapy*. Boston: Houghton Mifflin Company.

Rogers, C. R. (1961). *On becoming a person: A therapist's view of psychotherapy*. Boston: Houghton Mifflin Company.

Rule, W. R. (1991). Self actualization: A person in positive movement or simply an esteemed personality characteristic? *Journal of Social Behavior & Personality, 6*(5), 249-264.

Ryan, R. M. & Connell, J. P. (1989). 'Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology, 57,* 749-761.

Ryan, R. M., Huta, V., & Deci, E. L. (2008). Living well: A self-determination theory perspective on eudaimonia. *Journal of Happiness Studies, 9,* 139–170.

Ryff, C. D., & Keyes, C. L. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology, 69,* 719–727.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19(4)*, 460-475.

Sheldon, K. M. (2004). *Optimal human being: an integrated multi-level perspective.* Mahwah, NJ: Lawrence Erlbaum Associates.

Sheldon, K.M & Kasser, T. (1998). Pursuing personal goals: Skills enable progress, but not all progress is beneficial. *Personality and Social Psychology Bulletin, 24*, 1319–1331.

Sheldon, K. M., Kasser, T., Smith, K., & Share, T. (2002). Personal goals and psychological growth: Testing an intervention to enhance goal attainment and personality integration. *Journal of Personality, 70, 5-31.*

Sheldon, K. M., Ryan, R. M., Deci, E. L. & Kasser, T. (2004). The independent effects of goal contents and motives on well-being: It's both what you pursue and why you pursue it. *Personality and Social Psychology Bulletin, 30*, 475-486.

Shostrom, E. L. (1964). A test for the measurement of self-actualization. *Educational and Psychological Measurement, 24,* 207-218.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7,* 422–445.

Seligman, M. E. P. (2002). *Authentic happiness: Using the new positive psychology to realize your potential for lasting fulfillment.* New York, NY: Free Press.

Stevens, J. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Lawrence Erlbaum.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using Multivariate Statistics.* Boston, MA: Pearson Education. Inc.

Tarricone, P. (2011). *The taxonomy of metacognition*. New York: Psychology Press.

Tatarkiewicz, W. (1976). *Analysis of happiness.* The Hague: The Netherlands, Martinus Nijhoff.

Telfer, E. (1980). *Happiness.* New York, NY: St. Martin's Press.

Vella-Brodrick, D. A., Park, N., & Peterson, C. (2009). Three ways to be happy: Pleasure, engagement, and meaning—Findings from Australian and US samples. *Social Indicators Research, 90*(2), 165-179.

Waterman, A. S. (1990a). Personal expressiveness: Philosophical and psychological foundations. *The Journal of Mind and Behavior, 11* (1), 47-74.

Waterman, A. S. (1990b). The relevance of Aristotle's conception of eudaimonia for the psychological study of happiness. *Theoretical & Philosophical Psychology, 10*(1), 39-44.

Waterman, A. S. (1993). Two conceptions of happiness: Contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *Journal of Personality and Social Psychology*, *64*(4), 678-691.

Waterman, A. S., Schwartz S. J., Zamboanga B. L., Ravert R. D., Williams M., Agocha V. B., Kim, S. Y., Donnellan, M. B. (2010). The Questionnaire for Eudaimonic Well-Being: Psychometric properties, demographic comparisons, and evidence of validity. *The Journal of Positive Psychology*, *5* (1), 41-61.

Walker, B. A., Reis, S. M., & Leonard, J. S. (1992). A developmental investigation of the lives of women. *Gifted Child Quarterly, 36*,(4), 201-206.

Culture-related factors affect sunk cost bias

Carol Y. Yoder, Ruben Mancha and Nupur Agrawal

Trinity University

ABSTRACT

Reasoning and decision-making are fraught with systematic errors in thinking. One key example is sunk cost, a past investment that cannot be recovered, that influences ongoing decisions. A sunk cost bias occurs when previous choices affect present decisions. Sunk cost decision-making has been primarily studied in Western, individualistic cultures although some attention has been focused on comparing its prevalence in collectivist cultures such as Japan and China. We evaluated the influence of individualist and collectivist cultures, perceived control and the role of self. In Study 1 Americans and Indians were primed with cultural values and then presented with sunk cost decision scenarios. Results indicated Americans made more sunk cost decision errors than Indians and personal decisions were associated with more bias than decisions made on behalf of others. Cultural differences on sunk cost bias were consistent with self-justification theory. In Study 2 a new set of sunk cost scenarios varied environmental use and sustainability themes. Results indicated particular situations influenced error, although country of origin and perceived behavioral control were also effective at predicting sunk cost bias.

KEYWORDS: decision-making, cognitive bias, sunk cost, self-view, India, United States, environment

REASONING AND DECISION MAKING are fraught with systematic errors in thinking, which adversely affect behavior. Much of the research on cognitive bias has been conducted on Western cultures with some comparisons to Eastern cultures, particularly China and Japan. Our purpose was to explore how different factors influence reasoning biases, specifically, how people with individualist cultural values in the United States differ from people with collectivist cultural values in India. To our knowledge, this is the first cognitive bias study reporting American and Indian responses to environmental issues designed to trigger sunk cost bias. While comparative research has studied sunk cost bias in China, Japan, and Mexico (Chow, Harrison, Lindquest & Wu, 1997; Greer & Stephens, 2001; Kitayama, Markus, Matsumoto, & Norasakkunkit, 1997), India's population and particular demographics are distinctive in many respects and therefore deserve independent scrutiny. Sunk cost biases, concerning environmental themes such as water and energy conservation, were used in this study; this particular content is salient because of the heavy demands on our environment for sustaining large populations and the implications for policy.

» DEFINING SUNK COST BIAS

A sunk cost is a past cost that cannot be recovered. A sunk cost effect occurs when past investment of time, money or effort influence a present investment decision (Arkes & Blumer, 1985). That prior investment is often considered the motivating factor in the decision to continue or escalate an investment. A sunk cost choice (i.e., a choice taking into account a sunk cost) is considered irrational because decisions are based on past investments, rather than on unbiased future outcomes. This cognitive bias can be very costly across a wide range of resources (e.g., Ross & Staw, 1993).

» WHY SUNK COST BIAS OCCURS

In business, economics and psychology, sunk cost has been used to explain choices as diverse as the maintenance of troubled relationships to financial speculation (e.g., Strough, Schlosnagle, Karns, Lemaster & Pichayayothin, 2013). Sunk cost bias provides a plausible explanation for poor decisions regarding the space shuttle Challenger (Economist, 2003), any number of military incursions into other countries, or even why we persist with some

Author note: Carol Y. Yoder, Department of Psychology, Trinity University; Ruben Mancha, Department of Finance & Decision Sciences, Trinity University; Nupur Agrawal, Department of Psychology, Trinity University. Correspondence concerning this article should be addressed to Carol Yoder, Department of Psychology, Trinity University, San Antonio, TX 78212, cyoder@trinity.edu

of the energy choices we make (Muehlenbachs, 2013). Some sunk cost research has focused on individuals' rationales when making sunk cost choices. Arkes and Blumer (1985) determined that the psychological justification for making sunk cost choices has to do with the desire to avoid waste (see also Bornstein & Chapman, 1995). That is, people do not want to appear extravagant or inefficient. Learning lessons from poor past decisions has also been provided as a secondary explanation (Bornstein & Chapman, 1995). In explaining their reasoning, people want to believe they are not repeating previous mistakes.

Apart from avoiding the appearance of waste, research on sunk cost has often pointed to two theories addressing why this bias occurs: prospect theory and self-justification theory. Prospect theory explains how people evaluate outcomes in terms of gains and losses. Prospect theory describes actual behavioral actions, where people underweight outcomes that are probable in comparison with those that seem certain (Kahneman & Tversky, 1979). When a decision is framed as a loss or when previous investments are involved, we are more willing to take risks to avoid the loss (Thaler, 1980). This corresponds to Whyte's (1986) finding that people are more risk-averse in gain situations and more risk-prone in loss situations. For example, in a gain situation, people prefer a certain win of $50 over a fifty percent chance of winning $100 or $0, and in loss situations people would rather take a fifty percent chance of losing $100 or $0 than a certain loss of $50 (Brockner, 1992). The destabilizing tendencies of loss aversion increase people's risk tolerance and the likelihood of sunk cost bias (Whyte, 1986).

Alternatively, self-justification theory posits that people do not like to admit error (Brockner, 1992). In sunk cost situations, self-justification theory predicts that people will choose to continue with an investment to 'reaffirm' their original investment decision was 'correct' (Brockner, 1992). Staw (1976) found that when feedback was negative, and decisions involved personal responsibility, people were more likely to escalate investment. This is congruent with self-justification theory, because Brockner (1992) found that both negative feedback and personal responsibility encourage reaffirmation of the original choice. However, a related concern is maintaining one's reputation for completing actions (Dixit & Pindyck, 1994; Milgrom & Roberts, 1992), even in the case of expected loss. Self-justification theory is important to the present study because of its focus on the self. How the self is viewed depends on culture (Markus & Kitayama, 1991), and self-justification theory may contribute to understanding sunk cost choices.

» PERCEIVED RESPONSIBILITY AFFECTS SUNK COST CHOICES

Responsibility to a group of peers, as opposed to personal responsibility, also impacts sunk cost bias, but not in consistent ways. Staw and associates (1976) manipulated responsibility and found that more responsibility escalated commitment. Arkes and Blumer (1985) followed up, noting the role of personal responsibility. More recently, Wong (2005) also found having a personal stake in a decision increased the likelihood of escalating commitment.

Alternatively, Simonson and Nye (1992) found that heighted responsibility improved decision-making, such that participants in a higher responsibility condition were less likely to make sunk cost choices than those with lower responsibility. In other words, participants who thought they had to explain their decisions to others and/or be evaluated by others were less likely to make sunk cost choices than those who thought their decisions were confidential. Accountability in the form of monitoring also reduced escalation of commitment (Kirby & Davis, 1998). Institutional decision makers must be able to explain themselves. If they are accountable to others, they may be more motivated to attend to and process information (Simonson & Nye, 1992). Petty and Cacioppo (1986) suggested either a central or more peripheral path of managing information, with higher responsibility tasks recruiting more elaborative, effortful path processing (Johnson & Eagly, 1989). When making judgments for ourselves, we might be less motivated to effortfully process information, bolstering our choices with self-justification. When making judgments for others, we may be more inclined to think carefully.

» CULTURE

Culture can be a potent influence on behavior. Often characterized by the dimension of relative collectivism or individualism (Triandis, 1995) Hofstede (1984) explains this dimension by a given culture's relative focus on the self or their group. Individualists are focused on promoting, and doing what is best for themselves, while collectivists are focused on care of others and harmony of the group (Hofstede, 1983; Hofstede, 1984). Geiger, Robertson, and Irwin (1998) applied Hofstede's (1983) distinctions of culture to sunk cost situations using self-justification theory. In self-justification theory a key focus is the personal desire to be correct and to fortify one's choices (Brockner, 1992). Because of the focus on the self, people from individualist cultures might be more likely to make sunk cost errors; following self-justification theory, they might reify and reaffirm their past decisions (Geiger et al., 1998). In contrast, collectivists might commit less sunk cost bias because their decisions may be focused on optimizing outcomes for the group.

On the other hand some business-oriented work has found that Chinese participants are more likely to escalate their commitment to projects, relative to U. S. subjects (Chow et al., 1997). Chinese participants may be more concerned about saving face, resulting in more commitment to prior decisions; however, that may be balanced by a greater willingness to tolerate risk. In a similar vein, Sharp and Salter (1997) found that Asian managers made riskier decisions than North American managers, particularly when there were long-term potential benefits for their firms. However, this seems to depend on time trajectories. Asian managers were less risky relative to North Americans when making short-term financial decisions but also more willing to take a longer-term orientation towards problem solving when making decisions. In another study Khan, Salter and Sharp (2000) manipulated previous resource investment and responsibility and found their multinational sample comparing Pacific Rim countries with North America had different responses to these dimensions. While all respondents were sensitive to framing, risk tolerance and personal involvement differed based on economic expectations and short or long-term time projections.

Weber and Hsee (1998) also found that people in individualist cultures took fewer risks than people in collectivist cultures. When comparing decisions made by Chinese, American, German, and Polish students, Weber and Hsee reported that American participants were considerably more risk-averse than their Chinese counterparts. In explaining these results, they argued that the cultural mesh in collectivistic societies offers the Chinese a "cushion" that reduces their perception of danger. This different perspective may make them appear less risk-averse although that may be because they rely on other contingencies.

To recap, there are a variety of cultural influences on sunk cost with some research pointing to individualistic cultures being more susceptible to bias, due to self-justification. Alternatively, other research finds more sunk cost bias in collectivist samples, explained in part by the benefits of group protection or different cultural expectations. Not surprisingly, type of issue or content under consideration is probably salient.

» COMPARING CULTURALLY-PRIMED AMERICANS AND INDIANS ON SUNK COST BIAS

In the first study, we primed participants either congruently or incongruently with their cultural values before asking them to read and respond to vignettes designed to elicit sunk cost bias. The goal of the priming intervention was to amplify cultural salience strategically, using the logic of experimentation, to separate influences. We evaluated three hypotheses in this study. First, we hypothesized that, like Geiger et al. (1998) and Kitayama et al. (1997) suggest, individualists will show more sunk cost bias than collectivists (H1), because of their enhanced self-focus (Hofstede, 1983). As the second hypothesis we posited that individualists should be more prone to sunk cost errors under conditions of personal responsibility (H2) as opposed to institutional responsibility, where responsibility is to peers as well as the self. Being more aware and accountable to others might increase deliberative thinking. In contrast, type of responsibility should have little effect on collectivists' sunk cost choices because they routinely attend to others. By definition, collectivists show more self- criticism (Kitayama et al., 1997) and focus on the group (Hofstede, 1983). That perspective should lead to greater focus on what is best for others. Who is or is not responsible should not be salient.

Our third hypothesis was that culture-consistent priming should enhance any sunk cost bias effects (Oyserman & Lee, 2008), while culture-inconsistent priming might balance or negate typical responses. That is, culture-consistent priming might accentuate sunk-cost errors across conditions of personal and institutional responsibility (H3). Oyserman and Lee suggest that priming all participants reduces questions about inference since one cannot assume that non-primed participants do not have a cultural frame of mind. Nevertheless, they point out that the concept of a priming manipulation is decidedly Western and note problems of cross-cultural priming comparisons. Previous research has tried to enhance cultural salience through priming. Oyserman and Lee (2008) review and critique different techniques. As will be discussed in the procedure, we modified a pronoun-priming task from Brewer and Gardner (1996), which Oyserman and Lee's

meta-analyses indicated was moderately effective. Participants were asked to select pronouns out of a word matrix, rather than a paragraph, to reduce the amount of reading required.

» METHOD

Participants

The sample of 438 participants included 204 English-speaking Indian students ($M_{age\ male} = 19.74$, $SD = 1.19$; $M_{age\ female} = 19.52$, $SD = .9$) from colleges and professional schools in Gujarat, India 233 and 233 American students ($M_{age\ male} = 19.22$, $SD = 1.36$; $M_{age\ female} = 19.16$, $SD = 2.05$) from a small private university. The sample included 82 Indian females and 142 U. S. females. Indian participants were recruited by the third author, with the administrative support of the various schools in Gujarat. American students were recruited with administrative support from psychology and business departments. Following the Triandis (1995) classification, American students were used to represent an individualist culture whereas the Indian students represented a collectivist culture. Although we had an English-fluent Indian sample, for many Indians in our sample, English was not their first language. In part, to ensure full comprehension of our task, we limited our sample to those participants who correctly and quickly identified pronouns in the priming task. Accuracy on the pronoun-priming task indicated both understanding of the English language as well as motivation to be an engaged participant. Prior to applying this criteria, the original sample of 666 subjects included an additional 57 American and 171 Indian participants.

Materials and design

This study used a 2×2×2 ([country: India; U. S.] × [priming: culture-consistent vs. culture-inconsistent] × [responsibility: personal vs. institutional]) mixed design, with responsibility as a within-subjects factor. All materials used in this study were environmentally-themed and data collection used a web survey.

A modified version of Brewer and Gardner's (1996) and Gardner, Gabriel and Lee's (1999) priming task was used to increase the salience of cultural values. Participants viewed a matrix of 81 words in which 14 were target-pronouns; the rest were non-target nouns with environmental themes (e.g., air, oil). Participants were instructed to click the target pronouns as quickly as possible. Through computer-generated random assignment, half of the participants were presented with matrices containing only individualistic target pronouns (e.g., I, my, mine). Half were presented with matrices containing only collectivistic target pronouns (e.g., we, us, ours). The number of correct responses and completion times were recorded; participants with pronoun identification scores of less than 10 or longer response times (outliers from the normal distribution) were excluded from our sample to ensure comprehension and motivated participation. See Appendix A for this measure.

Upon completing the pronoun-priming task, participants read and made decisions using information presented in four sunk cost vignettes similar to stimuli used in past literature (Bornstein & Chapman, 1995; Navarro & Fantino, 2009; Rosenboim, Shavit, & Shoham, 2010). Each of the four vignettes described an environmental dilemma in the context of whether there was personal or

institutional responsibility, which was counterbalanced. Each situation framed a past investment decision as resulting in a likely loss and suggested alternative solutions. For example, participants read:

> "It is important to **you/your company** to behave responsibly toward the environment and you are well aware of the projections of future water shortages. After years of planning your new low environmental impact **home/facility**, construction has begun. In trying to get the project off the ground and take advantage of short-term cost savings due to the slow economy, **you/your company** decide to purchase supplies for a traditional plumbing system for the new construction. In addition, this decision requires a non-refundable payment that covers one-third of the total cost of installation."

After talking to a colleague who has been researching an innovative rainwater collection system, you learn that it would be feasible and more environmentally friendly to install this new system. This system does not allow **you/your company** to use the traditional plumbing supplies or the labor which has already been paid. Although this new system is not cheap, it is highly efficient without ongoing costs. However, admittedly it has not been tested over a substantial period of time. Should you continue with the partially paid traditional plumbing system or switch to the new rainwater collection system." See Appendix B for the remaining vignettes.

Vignette Response Format and Coding. Participants could choose to stay with the original plan, switch their course to the new plan, or present an alternative. This involved clicking on one of these choices; the alternative choice provided a box to type in a response. Responses ranged from 'stay,' coded as a 1, which indicated sunk cost bias, to 'switch,' coded as a 5. 'Alternative' responses ranged from 2 to 4, depending on their rationale, with 3 indicating a more novel alternative proposed by the subject. When their written alternative was highly similar but somewhat different from the described sunk cost, it was coded a 2. If their alternative was similar but at least somewhat different to the proposed new technique, it was coded a 4. A score of 3 reflected a more unusual or hybrid approach. Each vignette was scored separately. Lower scores indicated more sunk cost bias and preference for continuation of the status quo.

The first author created an extended coding scheme based on pilot data, with detailed explanations for determining a score of 2, 3, and 4. Judgments were only required for those respondents who selected 'alternative' and provided an explanation of their choice. As previously noted sunk cost bias was reflected with a response of stay and choosing the described alternative was coded as a 5. Four undergraduates were trained by the first author and then assigned to score alternative decisions from particular vignettes. Two raters, blind to respondent characteristics, coded all of the alternative choices from the four vignettes. Inter-class correlations were computed to determine inter-rater reliability; coefficients ranging from .72 to .93, denoting good to excellent agreement (Cicchetti, 1994).

Although individual scores were computed for each vignette, for our primary analyses we were more interested in our manipulated variables than in individual responses to particular vignettes, so we combined vignettes for each of the two responsibility conditions. Responses to the two institutional perspectives and the two personal perspectives were separately averaged for each participant, yielding one primary dependent variable for institutional responsibility and one for personal responsibility. In addition, each vignette was looked at separately, to identify separate patterns unique to specific environmental issues (e.g., energy, water conservation).

» PROCEDURE

After providing informed consent, data was collected using an online web application. Initially, participants responded to demographic questions. Then participants completed the pronoun priming task where they were randomly assigned to be primed in ways that were either consistent or inconsistent with their country of residence. Culturally consistent priming consisted of Americans being primed with individualistic values (e.g., targets included mine, me). Culturally inconsistent priming involved words consistent with collectivistic values (e.g., targets included ours, we). Time spent on task was recorded, as were the number of correct pronoun identifications. After the priming task, participants read and responded to four randomly presented but counterbalanced sunk cost vignettes. Participants were asked to decide to stay with the choice where resources had already been invested, switch to an effective, newer method, or select a third option that required an explanation of their solution. Although no time limits were imposed, time spent on task was monitored and recorded. After completing study procedures, participants were thanked for their participation and provided contact information for follow-up, if they were interested in learning about outcomes.

» RESULTS

Initially we verified there were a sufficient number of sunk cost errors made ($m = 29.08\%$) in response to our environmental scenarios. There were more sunk cost errors in the U.S. relative to the Indian group, but percentages differed by vignette. Grouping responses by gender and country and separating by vignette, percentage of error ranged from as low as 9% (for Indian females) to as high as 63% (for American females). See Table 1 for numbers and percentage of error by vignette.

Separate chi-square analyses were used to test whether country of origin differentially affected decision. Responses were dichotomized as reflecting sunk cost (1) or solution switching (2-5). It was necessary to collapse responses to meet chi-square assumptions

Table 1. sunk cost error and percent listed by vignette theme for study 1

	coal energy		green growth		recycling		plumbing	
	count	percent	count	percent	count	percent	count	percent
India female	7	9	20	24	24	29	15	18
India male	15	12	31	25	46	37	27	22
U.S. female	22	15	30	21	66	46	90	63
U.S. male	19	21	42	32	37	41	54	59

* Not all participants responded to every computer-presented vignette.

of number of observations per cell. All four vignettes yielded significant country of origin effects. With Coal Energy, the Pearson chi-square analysis yielded $\chi^2_1 = 5.23$, $p = .02$. With Recycling, the results were $\chi^2_1 = 5.9$, $p = .01$; with Green Growth, the chi-square results were χ^2_1 of 3.7, $p = .05$, and Plumbing yielded a χ^2_1 of 120.77, $p < .001$. Americans made substantially more sunk cost errors, while Indians showed as much as 40% more likelihood of choosing the non-biased alternative.

To test our first hypotheses that country of origin would affect sunk cost, we ran a 2×2×2 ([country: India; U. S.] × [priming: culture-consistent vs. culture-inconsistent] × [responsibility: personal vs. institutional]) repeated measures ANOVA, with country of origin and type of pronoun priming as between subjects factors, and responsibility as a within subjects factor. As hypothesized, there was a significant main effect of country of origin, $F(1, 376) = 60.47$, MSE = 119.67, $p < .001$, $\eta_p^2 = .14$, such that Americans were nmore likely to make sunk costs errors than Indians (H1).

Our second hypothesis was that individualists should be more prone to sunk cost errors under conditions of personal responsibility as opposed to institutional responsibility, where responsibility is to others. Because of collectivists' orientation to others, we hypothesized that the responsibility manipulation should have little effect on collectivists' choices. Results were consistent with this hypothesis as well. We found an interaction between responsibility × country of origin interaction, $F(1, 376) = 5.9$, MSE = 11.57, $p < .02$, $\eta_p^2 = .02$ (see Figure 1). There was also a significant main effect of responsibility, $F(1, 376) = 11.64$, MSE = 22.83, $p < .001$, $\eta_p^2 = .03$, such that personal responsibility yielded more sunk cost errors than institutional responsibility (H2). However this effect was mostly due to differences in the American sample.

Figure 1. averaged sunk cost choices with error bars
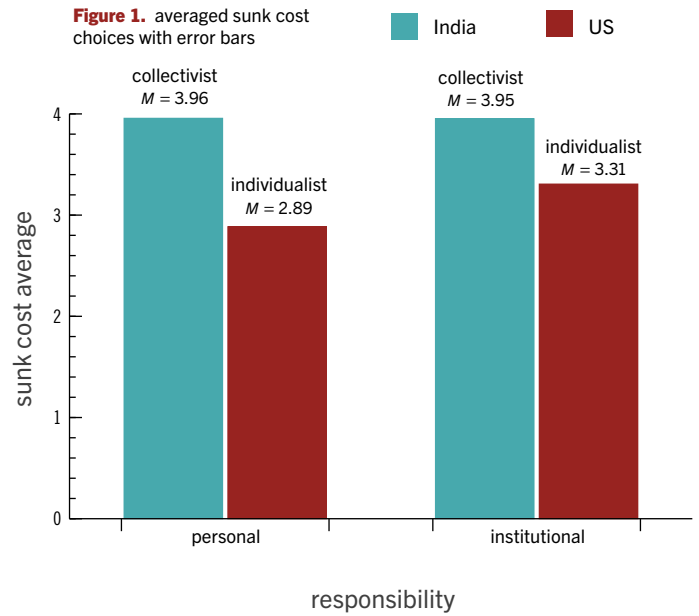
Our third hypothesis was that culture-consistent priming should enhance these effects. Although priming did not yield a significant main effect, there was a significant three-way interaction between priming × country of origin × responsibility, $F(1, 376) = 6.94$, MSE = 13.61, $p = .02$, $\eta_p^2 = .02$. See Figure 1. In analyzing this interaction, we found that scores for the Indian sample were more consistent than for Americans. Here, priming had little effect on the Indian sample. Among Americans, culture-consistent priming yielded sunk cost errors under conditions of personal responsibility (H3), as was predicted by the hypothesis. Priming

Figure 2. averaged sunk cost separated by priming, country and responsibility with error bars
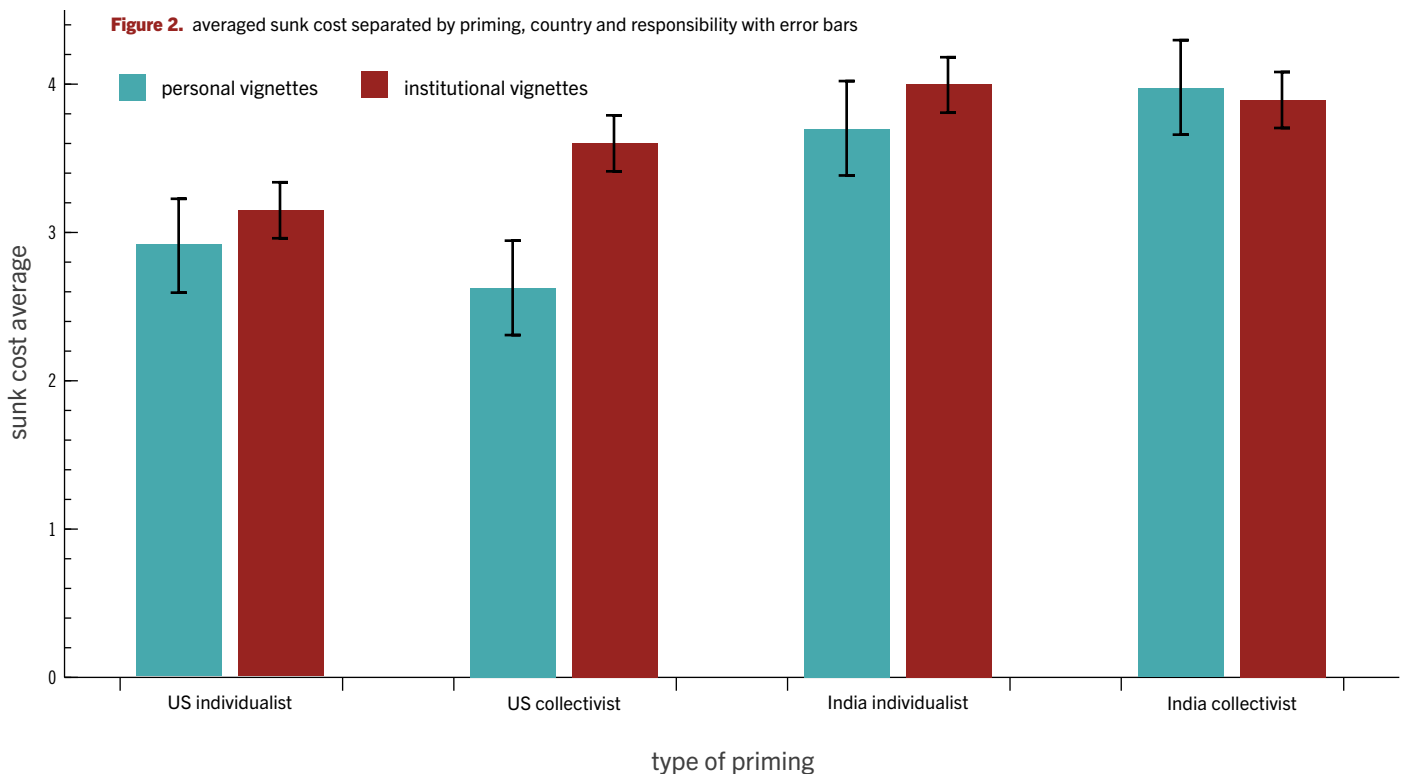
counter-to-culture for Americans (e.g., priming collectivist thinking) had a greater effect on decision-making by decreasing bias in institutional decisions, making Americans less likely to make sunk cost choices. For Americans, bias was similar in personal responsibility regardless of priming condition.

» DISCUSSION

All participants were more likely to choose alternatives rather than make sunk cost errors, but differences between countries, as well as within each country and within differing types of responsibility were present. There is no particular reason to hypothesize that perceptions of loss (or possible gain) should differ by culture as would be suggested by prospect theory. Nevertheless, this cannot be ruled out. The more straightforward explanation of our results provides support for Geiger et al.'s (1998) hypothesis that individualists would be more likely to make sunk cost choices than collectivists. It may be that individualists make more sunk cost errors because their focus on themselves encourages reaffirming their original decisions. Self-enhancement may be at the core (Kitayama et al., 1997), with greater ego involvement in choices, adversely affecting decision-making. Since self-report does not allow us to know the complete reasoning behind choices, perhaps Americans were more sensitive to minimizing loss as prospect theory suggests, or more attuned to avoiding waste.

Overall, participants from both countries made fewer sunk cost errors when presented with situations involving institutional responsibility. There was an interaction between country of origin and responsibility, such that Americans were more likely to make sunk cost errors in situations with personal responsibility. This interaction is also consistent with the self-justification hypothesis, because it suggests that sunk cost bias might be driven by an internal need to support one's previous decisions, especially for the American sample. Apparently, when decisions are made on behalf of others, previously committed resources play a less influential role relative to others' assessments.

Indians made significantly fewer sunk cost errors, but this could reflect more interest in innovative environmental technologies. Priming had less demonstrable effect on their thinking. In reflecting on these outcomes, one might infer that Indians were more comfortable with change. While this may be accurate, Hofstede (1983) studied the related concept of uncertainty tolerance, defined as a society's ability to deal with uncertainty and unstructured, ambiguous situations. In his 50-country comparison study Hofstede found that India and the United States were very comparable, with both accepting a moderate level of uncertainty in decision-making. Our data supports Geiger and associates' (1998) theorizing and our hypotheses.

Only modest support was found for our hypothesis that culture-consistent priming would enhance these effects. Priming consistent with culture did enhance Americans' bias in situations of personal responsibility. The three-way priming interaction suggested Americans were susceptible to thinking more collectively *only* when making institutional decisions when primed counter to culture. For Indians, the priming might not have been very effective because the Indian sample experienced trouble comprehending the pronouns. However, this did not seem to be an issue

because all included participants rapidly and correctly identified pronouns. As an additional test, we also compared an elite sample of only top performers (based on speed and complete pronoun accuracy) in this sample and still failed to get a priming main effect (H2). It is interesting that priming counter-to-culture appeared to increase Americans' awareness of others with institutional responsibility while Indians showed little differential response to the priming manipulation. Inconsistent priming in Americans with personal responsibility yielded outcomes statistically similar to consistent priming.

We were interested in trying to identify factors that might have influenced our outcomes in this first study. In Study 2, because of possible preference for green solutions more evident in Indians' decision making, we created vignettes where sunk cost bias could be examined apart from green or innovative technologies. In reflecting on our findings we recognized that choosing novelty and innovation over tradition is often confounded in sunk cost studies and this change represented that awareness. Recognizing that country of residence is a weak proxy for culture, it seems likely that some variation might be attributable to different self-views. While self-view is influenced by culture and daily experiences, it also affects how we interpret, conceptualize, and respond to problems and make decisions.

Construction of the self

How we think about our self begins quite early. By or before two years of age, the majority of American and Canadian children recognize themselves in a mirror (Broesch, Callaghan, Henrich, Murphy, & Rochat, 2010). In contrast, there were few spontaneous self-orienting behaviors in non-western children. Nevertheless, children begin to distinguish themselves from others in early childhood in ways that depend on cultural opportunities. It is unlikely these early differences disappear in adulthood.

Depending on your experience growing up, Kitayama et al. (1997) provided evidence that the self is construed as an "interdependent and mutually connected entity" in collectivist cultures and an "independent and autonomous entity" in individualist cultures. Research suggests collectivists are more self-critical, being more likely to accept the accuracy of negative feedback and less likely to rationalize that information (Kitayama et al., 1997; Heine & Lehman, 1997). Self-criticism can help identify shortcomings, in order to ameliorate deficiencies and better fit into cultural standards of appropriateness (Heine & Lehman, 1997; Kitayama et al., 1997). The effects of self-criticism should apply equally to personal decisions and institutional decisions because decision makers want to meet the high expectations of their family and business partners. Individualists, on the other hand, may justify suboptimal choices to maintain self-esteem (Kitayama et al., 1997), especially with more personal, ego-involving decisions.

When considering how self-concept might influence willingness to escalate commitment Sivanathan, Molden, Galinsky, and Ku (2008) suggests that self-affirmation could both increase and reduce escalation. Individuals show decreased escalation of commitment when self-esteem reaffirmations offered an alternative route to cope with failed decisions. On the contrary, when self-affirmation highlighted how one's abilities lead to the unsuccessful decision,

participants escalated commitment (Sivanathan et al., 2008). Living in a collectivist society may generate comfort in their peers' support and create feelings of general worthiness, which decrease susceptibility to sunk cost bias. However, regardless of one's cultural heritage, any given individual may or may not fully embrace prominent cultural values. Individuals may differentially value taking care of others or making an independent contribution or consider both of these orientations equally worthy. Originally suggested by Markus and Kitayama (1991), self-construal may capture and explain the effects of culture on a variety of behaviors. One's self image is conceptualized as encompassing how one thinks, feels, and acts in relation to others and oneself, as distinct from others (Singelis, 1994). Those who value interdependence are focused on fostering harmony with others and meeting their expectations. Those who highly embrace independence are governed more by their internal thoughts and feelings, with personal expression valued. Interdependence and independence are orthogonal (Singelis, 1994). While these characteristics may coexist in any individual person, self-construal focuses more on relative strength and influence. How people construe the self, in relation to others, could influence how we think about decisions, providing an explanation for differences between cultural groups.

How much control we perceive we have over our choices may also influence decision making, apart from designations of authority or responsibility. One construct that might differentiate choices we make is perceived behavioral control, which provides an estimate of ease or difficulty in accomplishing a behavior (Azjen & Madden, 1986). This involves beliefs about controllability and efficacy of behavior, although Terry and O'Leary (1995) suggest it is more about perception of external constraints. Cheng, Cheung, Chio, and Chan (2013) did a meta-analysis of 152 independent samples, representing adults from across the world. Lower sense of control was related to anxiety and depression, especially in individualist societies. They noted that less control does not have the same (negative) value across culture, but beliefs about control do influence thinking (Cheung et al., 2013), which may, in turn, influence whether or not sunk cost errors occurred.

In Study 2 we expected to replicate the previously identified country and responsibility effects (H1, H3). Additionally, we hypothesized that self-construal would influence likelihood of cognitive bias (H4). We also hypothesized that perceived behavior control would influence decision-making (H5).

» STUDY 2 METHOD

As in Study 1, environmentally-focused vignettes were created to measure decision making. Themes of environmental sustainability and use were separately represented, and there were two levels of responsibility (personal, institution) for all vignettes; the purpose was to extend previous results, predicting sunk cost error.

Sample

The sample consisted of 94 English-fluent university students from India (male M_{age} = 19.9; SD = 1.7; female M_{age} = 19.6; SD = 1.5) and 144 undergraduate college students (male M_{age} = 19.8; SD = 1.5; female M_{age} = 19.7; SD = 1.4) from the U.S. who agreed to voluntarily participate in the study. This included 40 Indian females and 80 American females. In order to ensure that the stimuli were fully comprehended, only participants who were fluent in English were included in the analysis. We excluded participants who scored less than 3 on the linguistic fluency item. Prior to applying this criteria, the original sample consisted of 103 participants from India and 159 subjects from the United States.

Materials and measures

English language skills. Participants provided information about their spoken languages, preferred language, and fluency with language. A 5-point Likert scale item was used to self-report fluency and comfort with English. Only participants who indicated they were comfortable and proficient with English were included in our analyses.

Sunk cost vignettes. We created four new vignettes containing information about environmental practices, which we labeled donor, carbon, recycle2, and water. To control for theme and particular content, vignettes were written such that the optimal decisions (i.e., no sunk cost error) required, in some cases, environmental sustainability choices or in other cases, environmentally unfriendly choices. In the former, optimal decisions were associated with pro-environment themes (e.g., minimal carbon footprint policy, donation to green NGO) while the environmental use vignettes were optimized by choices that required deciding not to use impractical and expensive "environmentally-friendly" options (e.g., costly recycling; unreliable rainwater collection system).

As in Study 1, we used two versions of each vignette. In the first version, the participant had to make a decision for just him or herself, thus engaging in personal responsibility. In the first version, the decision was made on behalf of a group, thereby engaging in institutional responsibility. Rather than the open-ended response used in Study 1, all participants were given only two choices, either to continue making their investment or to withdraw from that practice and engage in an alternative. Here, optimal choices were coded as 0 and sunk cost errors were coded as 1.

Self-construal (SC). Singelis (1994) characterized self-construal as an individual's self-representation, captured with two separate subscales measuring interdependence and independence. Interdependence was assessed through response to items such as "I will sacrifice my self-interest for the group I am in," while independence was measured with items such as "I enjoy being unique and different from others in many respects." Each subscale consists of 12 items, which participants responded to using a 7-point Likert scale from strongly agree to strongly disagree. See Cross, Hardin and Gercek-Swing (2011) for a recent review of this much-used measure.

Table 2. sunk cost error and percent listed by vignette theme for study 2

	donor		carbon		recycle 2		water	
	count	percent	count	percent	count	percent	count	percent
India female	11	31	17	47	31	89	20	57
India male	18	40	12	27	37	82	19	42
U.S. female	54	68	38	48	59	74	37	46
U.S. male	38	61	24	39	43	69	28	44

Table 3. logistic regression predicting sunk cost error on the donation to green practice (conservation).

variable	B	SE	p	odds ratio
country	1.28	0.32	0.00	3.59
responsibility	0.21	0.31	0.01	2.30
PBC	−0.05	0.04	0.19	0.96
SC-interdependence	−0.01	0.02	0.58	0.99
SC-independence	0.00	0.02	0.80	1.00
constant	−0.75	1.49	0.61	0.47

adjusted R^2 = 0.144

Perceived behavioral control (PBC). This 10-item measure was created by Kraft, Rise, Sutton and Roysamb (2005) and modeled after Azjen and Madden's (1986) estimate of perceptions about personal control and efficacy in accomplishing targeted environmental behaviors. This is an example item: "I am in full control of my actions to protect the environment." See Appendix D for this ten-item measure.

» PROCEDURE

After providing informed consent and responding to demographic questions in our online survey format, participants were presented with a series of four randomly presented vignettes. Since differential level of responsibility might affect propensity to sunk costs, the study was balanced such that each participant responded to two institutional and two personal vignettes as in Study 1. Similarly, each participant responded to two environment-conservation themes and two environmental-use themes. After providing their response to the vignettes, participants responded to the SC measure, and the PBC. All responses were collected online with Qualtrics software.

» RESULTS

Sunk cost errors were generally higher in Study 2 relative to Study 1, ranging from as low as 27% (Indian men) to as high as 89% (Indian women) when looking at each vignette separately (see Table 2 for numbers and percentage of error by vignette). In conducting a 2×2 (country × response) Pearson chi-square analysis, "Recycle 2" yielded χ^2_1 = 4.95, p = .03 and "Donor" yielded a χ^2_1 = 17.46, p < .001, indicating that rate of error systematically differed between respondents from India and the United States. Error rates were higher for U. S. participants on these two vignettes. Neither "Carbon" nor "Water" yielded significant differences in this basic analysis. In

Table 4. logistic regression predicting sunk cost error on the carbon footprint (conservation).

variable	B	SE	p	odds ratio
country	0.52	0.32	0.09	1.68
responsibility	0.02	0.30	0.96	1.02
PBC	−0.10	0.04	0.01	0.91
SC-interdependence	0.01	0.02	0.54	1.01
SC-independence	0.02	0.02	0.33	1.02
constant	−0.10	1.47	0.95	0.91

adjusted R^2 = 0.076

general it appeared that situations, which required respondents to make decisions that ran counter to environmental protection, elicited more error from Indians relative to Indians' decisions on vignettes that were more consistent with environmental protection. To try to capture these complexities, we compared the American and Indian participants on PBC, Self-Construal, and demographics and gender or country differences were found, indicating these groups were equivalent on these measures.

After assessing for multicollinearity and confirming that our measures were independent, logistic regression was conducted to assess whether country of origin, vignette responsibility type, PBC, SC interdependence and SC independence significantly predicted whether or not a participant made a sunk cost error. When all five predictor variables were considered together, they significantly predicted whether or not a participant made a sunk cost error. We found these results for all four environmental vignettes. Because our dependent variable of sunk cost error was dichotomous and particular environmental content influenced judgment, each will be presented separately.

Environmental conservation-oriented stimuli. When optimal decisions were oriented towards conservation, 114 participants committed the sunk cost error on 'donor' while 91 did not. Using all five predictors, the omnibus test of the model was significant, χ^2_5 = 23.30, p < .001, N = 205. Nagelkerke's R^2 estimated variance accounted for 14%. Of the 114 participants who did commit the sunk cost error, 81% were predicted correctly with this model but only 51% of the 91 participants who did not commit the error were correctly predicted. Thus overall, the model correctly predicts 67% of the participants.

U. S. participants had a 3.6 to 1 increase in the odds of committing a sunk cost error. Also, the odds of making an error were 2.3 times larger with personal responsibility, as opposed to institutional responsibility. That is, Americans were more likely to commit sunk cost bias and errors were more likely when participants made decisions for themselves as compared to when they made decisions on behalf of a group or organization.

The carbon footprint vignette was also written such that the optimal solution involved a pro-sustainability choice. The model significantly predicted error, χ^2_5 = 11.79, p = .038, n = 205, although only 8% of the variance was explained. Perceived behavioral control independently contributed to this model although there was a trend for Americans to make more sunk cost error, B = .52, p = .10.

One hundred twenty three participants did not commit sunk cost error while 82 did. Thus, if we were to consider participants who did not commit sunk cost error for this vignette, we would be correct 84% of the time. For the 82 that did commit sunk cost bias, only 29% were correctly predicted. Overall, the model correctly predicted choices in 62% participants.

Environmental use stimuli. With optimal decision making more oriented to using resources, errors on the recycle vignette were predicted by country and perceived behavioral control, χ^2_5 = 28.79, p < .001, N = 205.

This model was effective at predicting sunk cost error in 96% of cases but substantially worse at predicting who would not make this error (25%). Overall, this model accounted for 20% of the variance. The odds ratio of Country suggests that the odds of

making an error had roughly a 3:1 ratio for Indians. That is, with this environmental-use themed content, Indians were more likely to make sunk cost errors by choosing the greener but less optimal choice. The odds ratio of PBC is significant and indicates that for each point increase in perceived behavioral control, a subject has 87% of the odds of making a sunk cost bias. This suggests that being Indian and having less behavioral control contributes to a greater chance of making a sunk cost error.

In 'water', the second decision making situation that required an environmental use choice in order to avoid error, the omnibus test of the model was also significant, $\chi^2_5 = 10.99$, $p = .05$, $N = 205$. The model only accounted for 7% of the variance. Here, only perceived control separately predicted error. Once again, less perceived behavioral control was predictive of greater likelihood of sunk cost error. Country was not a separate predictor of bias.

For this vignette, 110 participants made the logically correct choice whereas 95 made an error. Of the 110 that did not commit a sunk cost error, 68.2% were correctly predicted while only 42.1% of those who made errors were correctly predicted by this model. Overall, the model correctly predicted 56% of the errors made by the participants.

» DISCUSSION

Our respondents were strongly influenced by the described situations. Country of origin was not a consistent predictor of error, although Indians generally showed less cognitive bias, when choices were supportive of environmental sustainability. On decisions which required making choices that might seem environmentally unfriendly (recycle 2, water), people who perceived they had less behavioral control were more susceptible to cognitive bias. Perhaps, less perceived control elicited a greater need to be green or make a more socially acceptable choice in spite of contra-indicating factors. People who felt they had more control did not fall into this trap.

With the more typically framed situations, where avoiding sunk costs involve more environmentally friendly outcomes, American respondents were more vulnerable to error. Perceived responsibility was important in green choices, but not salient with regards to monetary donations. It appears that environmentally-themed circumstances influenced choices although some vignettes were easier to separate out ineffective previous costs relative to new circumstances. In particular, all groups found the impractical and expensive recycling system sufficiently attractive to reinvest in it, in spite of the fact that it was not economically or functionally viable. Participants, especially Indians, were more susceptible to error with this ineffectual recycling venture, perhaps reflecting different environmental experiences, stronger interest in making a sustainability choice, and/or less concern about inefficiency and cost. Desire to be ecologically appropriate may interfere with optimizing decision making, at least in young adult samples.

» OVERALL DISCUSSION

Likelihood of making sunk cost error is affected by many environmental circumstances. Values associated with culture play a role but our data indicate particular situations are sometimes more important. Country of origin contributed to explaining cognitive bias as did perceived behavioral control. Sunk cost choices were

Table 5. logistic regression predicting sunk cost error on the recycle 2

variable	B	SE	p	odds ratio
country	1.04	0.43	0.015	2.84
responsibility	0.40	0.36	0.27	1.48
PBC	−0.17	0.04	0.00	0.87
SC-interdependence	−0.02	0.02	0.33	0.98
SC-independence	−0.01	0.02	0.59	0.99
constant	3.09	1.77	0.08	22.02

adjusted $R^2 = 0.196$

more likely in individual situations and less likely when making a decision on behalf of others, especially for Americans. However, these relationships were small, explaining little of the variability in people's decision making.

As initially theorized by Geiger et al. (1998) and reported by Kitayama et al. (1997) Americans were more susceptible to sunk cost bias in six different environmental decisions. The two environmental- use vignettes failed to find this relationship; both vignettes commanded high error rates from all groups. Indians were equally likely to make a sunk cost error on recycling, where cost and quality were not fully considered and previous choices were reified. We do not have a way to determine how much this error reflects environmental experiences. The water vignette also yielded a high overall error rate (47%) that country of residence did not discriminate. Perhaps circumstances concerning water use generate enough concern for both the peoples living in the southwestern United States and India, who share environmental concerns regarding drought and water scarcity. People may be willing to grasp at straws to avoid further loss. As prospect theory suggests, loss of a scarce resource may promote less considered reasoning, especially when respondents perceive they have little control.

Based on Study 1 where sunk cost errors differed by country of origin, we hypothesized that self-construal would partially explain differences in error rates. Particularly we expected that high levels of individualism without the moderating influence of high levels of interdependence would be related to greater likelihood of cognitive bias. As reported, it did not independently contribute to predicting sunk cost error. While the construct of self-construal continues to be considered highly relevant to understanding cultural differences (Cross, Hardin & Gercek-Swing, 2011; Dean & Gardner, 2014), Levine et al.'s (2003) review and analyses raise concerns about the reliability of this construct. Even though self-construal was

Table 6. logistic regression predicting sunk cost error on the water

variable	B	SE	p	odds ratio
country	1.14	0.31	0.65	2.15
responsibility	0.21	0.29	0.47	1.23
PBC	−0.10	0.03	0.00	0.91
SC-interdependence	−0.01	0.02	0.50	0.99
SC-independence	−0.00	0.02	0.79	1.00
constant	2.96	1.45	0.04	19.18

adjusted $R^2 = 0.070$

not individually effective in predicting sunk cost bias, it is interesting that when Americans were primed counter- to-culture (in ways that highlighted others), they thought more carefully when representing others. This increased awareness of others appeared to decrease their cognitive bias. This evidence does indirectly support the notion that self-view influences Americans' decisions and perhaps contributes to sunk cost error.

Perceptions of control were predictive of choices in both studies. In Study 1, when people made decisions on behalf of the group, fewer errors occurred. In Study 2, the institutional focus was more neutral and less personally involving than the vignettes in Study 1. This manipulation had less effect but the individual-level self-report of perceived behavioral control did predict responses in three of the four vignettes. People with a stronger sense of control of their behavior were less susceptible to making sunk cost errors. The sole vignette (donor) that was not affected by PBC involved a monetary donation that may have been less effective because the finances were hypothetical and the action more passive. It is interesting that personal and institutional responsibility was an effective predictor in Study 1, but not in any other vignette in Study 2 where PBC was significant. Whether we manipulate control or we measure people's behavioral intent, both of these efforts capture perceptions of control. Our findings from Study 2 suggest that perceived behavioral control is the more relevant construct (in contrast to manipulated personal and institutional responsibility). Not surprisingly, personal intent with regards to choices is more salient than responsibility, at least as manipulated in our vignette scenarios.

It is important to recognize beliefs and attitudes impact the assumptions that influence our choices, and these are influenced to some extent by culturally-influenced experiences. While we may ostensibly value careful reasoning, Americans and perhaps individualists more generally, may be unduly influenced by our need to protect both the self and previous choices, as self-justification theory suggests. Nevertheless, self-construal did not separately capture the group differences we found, although the priming interaction, which decreased American sunk cost error, does provide indirect support for self-focus contributing to cognitive bias.

Most of our effects accounted for little of the total variance. This may be partly due to our focus on environmental themes. While an environmental focus may be universally applicable, it may also decrease individual sense of responsibility and control to a greater extent than other issues. Because of the nature of environmental problems and the relatively small impact that one individual can have, even when acting on behalf of an entire corporation, one's choices may seem small when compared to the global scale of the issue. The diminished effectiveness of one person may have led to de-individuation effects that opposed the effects of responsibility to the self and to others. However, it is interesting that personal control mattered in predicting cognitive bias, in the different approaches of control used in Study 1 and Study 2.

Cultural equivalence is another concern, as is our exclusive use of English. Our sample was relatively young, and either engaged in their college years or slightly older; more heterogeneity in age would have likely increased response variability with people having more varied life experiences.

Given the identified cultural differences in sunk cost bias, our results suggest that different strategies for training and persuasion could be effectively used to educate, garner support for policy, or promote better decision making. It may also be the case that Indians are more comfortable and accustomed to change, and hence are more insulated from making the error of continuing to pour resources into an unsuccessful venture because of past investment. If Americans and those who are high in independence initially 'buy in' and do so without accountability to others, they may feel increased commitment to continued investment in further related decision making. Continuing investments following sunk costs may reflect living with past decisions, however suboptimal, because one is ultimately on one's own in individualist cultures. For individualists, caution is required; there is no safety net (Weber & Hsee, 1998). Reminding individualists of their responsibility to others may generate a more thoughtful analysis. When aware that others will make separate assessments of one's choices, individualists may apply more effortful strategies that involve more integrative and critical thinking. People with more interdependent values may be less likely to commit additional resources to failing ventures, at least when perceptions of risk are similar. Encouraging people to recognize their responsibility and agentic role also should improve decision making, since having a stronger sense of control was related to better choices.

Since we compared participants with two countries exuding high verticality, perhaps that dimension requires more scrutiny. Keil et al. (2000) also makes a convincing case for the role of uncertainty avoidance and suggests that group factors such as cohesion may interact with individual factors.

Altogether, this research points to factors that affect sunk cost choices. At least with sustainability situations, Indians made fewer sunk cost choices and Americans were more susceptible to errors, especially when making personal decisions. Likelihood of sunk cost error was greater for all respondents where optimal decisions involved leaving behind events previously considered environmentally innovative or effective. Interest in technology or innovation, environmental concern, and social desirability may also contribute.

With increasing communication and interaction between different cultures, it is important to know how different values can affect our reasoning and decisions. More research on cognition and common errors in thinking will improve our understanding of differences and promote better methods of working together. This seems particularly important with the current rate of globalization and recent increases in informational exchange that frequently function under different assumptions, modes of operation, and internal theories of appropriate behavior. ■

REFERENCES

Arkes, H.R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, *35*(1), 124-140.

Ajzen, I., & Madden, T.J. (1986). Prediction of goal-directed behavior: Attitudes, intentions and perceived behavioral control. *Journal of Experimental Social Psychology, 22*, 453-474.

Bornstein, B.H., & Chapman, G.B. (1995). Learning lessons from sunk costs. *Journal of Experimental Psychology: Applied*, *1*(4), 251-269.

Brewer, M.B., & Gardner, W. (1996). Who is this "we?" Levels of collective identity and self representations. *Journal of Personality and Social Psychology, 71*(1), 89-93.

Brockner, J. (1992). The escalation of commitment to a failing course of action: Toward theoretical progress. *Academy of Management Review, 17*(1), 39-61.

Broesch, T., Callaghan, T., Henrich, J., Murphy, C. & Rochat, P. (2011). Cultural variations in children's mirror self-recognition. *Journal of Cross-Cultural Psychology, 42*, 1018-1029.

Cheng, C., Cheung, S.F., Chio, J.H., & Chan, M.P. (2013). Cultural meaning of perceived control. A meta-analysis of locus of control and psychological symptoms across 18 cultural regions. Psychological Bulletin, 139 (1), 152-188. Doi: 10.1037/a0028596.

Chow, C.W., Harrison, P., Lindquist, T. & Wu, A. (1997). Escalating commitment to unprofitable projects: Replication and cross-cultural extension. *Management Accounting Research, 8*, 347-361.

Cicchetti, D.V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284-290.

Cross, S.E., Hardin, E. & Gercek-Swing, B. (2011). The what, how, why, and where of self-construal. *Personality and Social Psychology Review, 15*, 142-179.

Dean, K. & Gardner, W. (2014). How will "I" versus "We" perform? An Investigation of future outlooks and self-construals. *Personality & Social Psychology Bulletin, 40*, 947-958.

Dixit, A. & Pindyck, R.H. (1994). *Investment under uncertainty.* Princeton, NJ*:* Princeton University Press.

Economist (2003). The space shuttle: Old, unsafe and costly. *Economist* Magazine, Aug. 28, 2003, 77.

Gardner, W.L., Gabriel, S. & Lee, A. (1999). "I" value freedom but "we" value relationships: Self-construal priming mirrors cultural differences in judgment. *Psychological Science, 10*, 321-326.

Geiger, S.W., Robertson, C.J., & Irwin, J.G. (1998). The impact of cultural values on escalation of commitment. *International Journal of Organizational Analysis*, *6*(2), 165-176.

Greer, C.R. & Stephens, G.K. (2001). Escalation of commitment: A Comparison of differences between Mexican and U.S. decision-makers. *Journal of Management, 27,* 51-78.

Heine, S.J. & Lehman, D.R. (1997). The cultural construction of self-enhancement: An examination of group-serving biases. *Journal of Personality & Social Psychology, 72,* 1268-1283.

Hofstede, G. (1983). Dimensions of national cultures in fifty countries and three regions". In J.B. Deregowski, S. Dziurawiec and R.C. Annis (Eds), *Expiscations in cross-cultural psychology.* Lisse NL., Swets and Zeitlinger, 335-355.

Hofstede, G. (1984). *Culture's consequences*. Thousand Oaks: Sage Publications.

Johnson, B. & Eagly, A. (1989). Effects of involvement on persuasion: A Meta-analysis. *Psychological Bulletin, 106*, 290-314.

Kahneman, D. & Tversky, A. (1979). Prospect theory: An Analysis of decision under risk. *Econometrica, 47*, 263-291.

Keil, M., Tan, B.C.Y., Wei, K., Saarinen, T., Tuurnainen, V., & Wassenaar, A. (2000). A Cross-cultural study on escalation of commitment behavior in software projects. *MIS Quarterly, 24*, 299-325.

Khan, B., Salter, S. & Sharp. D. (2000). Pouring good money after bad: A Comparison of Asian and developed country managerial decision-making in sunk cost situations in financial institutions. *Asia-Pacific Development Journal, 7*, 105-120.

Kirby, S. & Davis, M. (1998). A study of escalating commitment in principal-agent relationships: Effects of monitoring and personal responsibility. *Journal of Applied Psychology, 83*(2), 206-217.

Kitayama S., Markus, H.R., Matsumoto, H., & Norasakkunkit, V. (1997). Individual and collective processes in the construction of the self: Self-enhancement in the United States and self-criticism in Japan. *Journal of Personality and Social Psychology, 72*(6), 1245-1267.

Kraft, P., Rise, J., Sutton, S. & Roysamb, E. (2005). Perceived difficulty in the theory of planned behavior: Perceived behavioral control or affective attitude? *British Journal of Social Psychology, 44*, 479-496.

Levine, T.R., Bresnahan, M.J., Park, H.S., Lapinski, M.K., Wittenbaum, G., Shearman, S.M., Lee, S.Y., Chung, D. & Ohashi, R. (2003). Self-construal scales lack validity. *Human Communication Research, 29*, 210-252.

Markus, H.R. & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*, 224-253.

Milgrom, P. and J. Roberts (1992). *Economics, organization, and management.* Englewood Cliffs, NJ: Prentice-Hall.

Muehlenbachs, L.A. (2013). A dynamic model of cleanup: Estimating sunk costs in oil and gas production, *RFF Discussion Paper 12-12-REV*. Washington, DC: RFF Press, Routledge.

Navarro, A.D. & Fantino, E. (2008). The sunk-time effect: An exploration. *Journal of Behavioral Decision Making*, *22*(3), 252-270.

Oyserman, D., & Lee, S.S. (2008). Does culture influence what and how we think? Effects of priming individualism and collectivism. *Psychological Bulletin*, *134*(2), 311-342.

Petty, R.E. & Cacioppo, J.T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology (Vol. 19*, pp. 123-205). Orlando, FL: Academic.

Rosenboim, M., Shavit, T., & Shoham, A. (2010). Financial decision making in collective society—A field test on Israeli kibbutz members and city residents. *The Journal of Socio-Economics*, *39*(1), 30-36.

Ross, J. & Staw, B.M. (1993). Organizational escalation and exit: Lessons from the Shoreham nuclear power plant. *Academy of Management Journal, 36,* 701-732.

Sharp, D.J. & Salter, S.B. (1997). Project escalation and sunk cost: A Test of the international generalizability of agency and prospect theories. *Journal of International Business Studies*, *28*, 101-121.

Simonson, I. & Nye, P. (1992). The effect of accountability on susceptibility to decision errors. *Organizational Behavior and Human Decision Processes*, *51*(3), 416-446.

Singelis, T.M. (1994). The measurement of independent and interdependent self-construals. Personality and Social Psychology Bulletin, 20(5), 580-591.

Sivanathan, N., Molden, D., Galinsky, A., & Ku, G. (2008). The promise and peril of self-affirmation in de-escalation of commitment. *Organizational Behavior and Human Decision Processes, 107*, 1-14.

Staw, B.M. (1976). Knee-deep in the big muddy: A Study of escalating commitment to a chosen course of action. *Organizational Behavior & Human Decision Processes, 16*(1), 27 − 44.

Strough, J., Schlosnagle, L., Karns, T., Lemaster, P., & Pichayayothin, N. (2013). No time to waste: Restricting life-span temporal horizons decreases the sunk-cost fallacy. *Journal of Behavioral Decision Making.* DOI: 10,1002/bdm.1781.

Terry, D.J. & O'Leary, J.E. (1995). The theory of planned behaviour: The effects of perceived behavioural control and self-efficacy. *British Journal of Social Psychology, 34*, 199-220.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization,* 1, 39-60.

Triandis, H. (1995). Individualism & collectivism. Boulder, CO US: Westview Press.

Weber, E., & Hsee, C. (1998). Cross-cultural differences in risk perception, but cross-cultural similarities in attitudes toward perceived risk. *Management Science, 44*(9), 1205-1217.

Whyte, G. (1986). Escalating commitment to a course of action: A Reinterpretation. *Academy of Management Review*, *11*(2), 311-321.

Wong, K. (2005). The role of risk in making decisions under escalation. *Applied Psychology: An International Review, 54*(4), 584-607.

APPENDIX A

›› CULTURE PRIMING WITH CULTURE-CONGRUENT & INCONGRUENT PRONOUNS*

Instructions. In this task, you will be presented with a matrix of words. The task is to select the pronouns (e.g., we, us) as quickly and accurately as possible by clicking once on the word. Your time to completion will be recorded.

Tree wildlife soil desert ecology *mine/we* forest reserve glacier

Snow rock tiger koppe *me/us* mist crust cactus *my/our*

Me/us tundra earth swamp typhoon cyclone *I/we* oil dust

Air ice *myself/ourselves* water crops core dam mineral rain

Park *mine/we* plains *I/ours* grass weather hawk resource crater

Cloud bear acidity stream river valley *mine/ours* power sand

Myself/we lake habitat crystal lava bee vapor volcano force

Cycle fern ocean fog *me/we* drought carbon *mine/our* drizzle

Sea nature terrain erosion *me/us* energy flood *I/we* climate

*Either individualist (I, me, my) OR collectivist (us, our, we) pronouns were displayed in regular non-boldfaced font.

APPENDIX B

›› SUNK COST VIGNETTES* FOR STUDY 1

Please read and select the option that best fits your perspective. You may describe your thoughts in the alternative response box.

Coal energy

The government has been pressuring Alero International to make business practices more environmentally conservative. Two years ago *you as Chief Financial Officer/The Board of Alero* decided to refurbish its coal burning machines at significant cost. The equipment is under an extended warranty for another eight years. In addition, four years remain on the contract with Alero's coal supplier.

A new system of machines has been introduced on the market using hydropower, which is said to be more cost effective. If Alero changes to these new machines, factory emissions would be reduced by 40% and production and labor cost might decrease as much as 15%. However, the initial cost of these machines is substantial and would require extensive retraining of staff. Although a small portion of these costs would be offset by government tax rebates, installation would result in revenue loss due to construction and employee training. In addition, contract buyouts are unpredictable, but could be costly. If this equipment update should somehow fail, Alero might not survive. On the other hand, success would bring much-needed positive attention to *you as their manager/Alero*. Should *you/Alero* stick with the coal burning equipment or change to the hydro-powered machines?

A) Stay with coal machines

B) Move to hydro-power machines

C) Alternative (explain)

Green growth

Green Growth, Inc. openly advertises their commitment to green policies and products. In your role as *chief financial officer/newly hired associate*, you/your company are/is contemplating buying a new car/fleet of cars to send an important message about you and Green Growth's commitment to sustainability.

You/your company currently *manage/owns* a four door sedan/

fleet of sedans purchased three years ago that will be paid off in full next year. The new hybrid car/fleet you were considering will potentially decrease gas cost substantially and will help promote Green Growth's core values. Trading in your current *fleet/car* will only cover 20% of the cost of the new *fleet/car* and will require you/your company to agree to a 60 month financing plan. This substantial cost will not allow you to invest in resources as you had planned. However, practicing what you preach builds credibility. Do you buy the new hybrid *fleet/car* or stick with the current transportation?

A) Stay with older transportation

B) Go with the hybrid

C) Alternative (explain)

Recycle

Support and costs for recycling in your city have risen steadily in the last three years. *As a city manager you must/Your city has decided to allow residents to* choose between continuing to use the same recycling group that has been in charge for the past 10 years, or to try a new system that advertises 33% lower monthly fees and reports use of more categories in sorting and better revenue from sold, recycled byproducts.

In reality, the old system worked fairly well and sorting was easy. The new system would require learning new sorting techniques. Additionally, *the city/you* would have to purchase new and more containers and this cost would be substantial. The new company guarantees the *city contract/your monthly rate* will be locked in for five years if the *city/you* decides to go with them. Do you stay with the old company or try out the new company?

A) Stick with the old company

B) Switch to the new company

C) Alternative (explain)

*Respondents saw either the institutional responsibility or the personal responsibility versions of these vignettes.

APPENDIX C

» PERSONAL AND INSTITUTIONAL SUNK COST VIGNETTES FOR STUDY 2

Environmental use

Please make a decision for the following problem: You have built an environmentally friendly home. The green materials and construction techniques increased the cost of your home by one third. Two months after moving in, a large storm ruined your environmentally friendly irrigation system using rainwater collection and solar energy. Should you rebuild the environmentally friendly system despite the high risk of a future storm damaging it, or should you opt for a standard irrigation system less sensitive to damage?

OR. Your company has built an environmentally friendly facility. The green materials and construction techniques increased the cost of your facility by one third. Two months after moving in, a large storm ruined your environmentally friendly irrigation system using rainwater collection and solar energy. Should your company rebuild the environmentally friendly system despite the high risk of a future storm damaging it, or should your company opt for a standard irrigation system less sensitive to damage?

A) Rebuild the environmentally friendly system despite the high risk of a future storm damaging it

B) Opt for a standard irrigation system less sensitive to damage

Please make a decision for the following problem. For five years you have been using an expensive service to recycle paper and plastic waste at a substantial cost each year. Because of your sustainable behavior, you have received several sustainability awards from your county. After relocating to a new city with fewer recycling initiatives, you could only find a recycling contractor of uncertain reputation that would charge you 100% more per year than the previous provider. Should you use the recycling service despite the cost increase or should you dispose of paper and plastic in the trash?

OR For five years, your company has been using an expensive service to recycle paper and plastic waste at a substantial cost each year. Because of your sustainable behavior, your company has received several awards from your local industry association. After relocating to a new city with fewer recycling initiatives, your company could only find a recycling contractor of uncertain reputation that would charge your company 100% more per year than the previous provider. Should your company use the recycling service despite the cost increase or should your company dispose of paper and plastic in the trash?

A) Use the recycling services despite the cost increase

B) Dispose of paper and plastic in the trash

Environmental conservation

Please make a decision for the following problem. You have decided to donate your supplemental pay this year to an NGO (Non-Governmental Organization) dedicated to the protection of the environment and global warming research. Before you complete the donation, your bank calls you to notify that, for the past five years, your account was subject to fraudulent charges adding up to more than your supplemental pay, which you had not noticed. Should you donate the money or not donate it and reconsider pledging next year?

OR: Your company has decided to donate the equivalent to all employees' supplemental pay to an NGO (Non-Governmental Organization) dedicated to the protection of the environment and global warming research. Before your company completes the donation, the bank calls to notify your company that, for the past five years, your company's account was subject to small fraudulent charges adding up to more than the annual supplemental pay, which your company had not noticed. Should your company donate the money or not donate it and reconsider pledging next year?

A) Donate the pledged amount

B) Not donate and reconsider pledging next year

Please make a decision for the following problem. You were planning on upgrading your car with a hybrid vehicle, paying an additional 40% to reduce your carbon footprint. A few days before completing the purchase, you learned that an investment had failed and you lost a similar amount of money. Should you buy the hybrid car or a gasoline car?

OR Your company was planning on upgrading its fleet of five cars with hybrid vehicles, paying an additional 40% (per car) to reduce your company's carbon footprint. A few days before completing the purchase, your company learned that an investment had failed and your company lost a similar amount of money. Should your company buy the hybrid cars or a fleet of gasoline cars?

A) Buy the hybrid car(s)

B) Buy the gasoline car(s)

APPENDIX D

» GREEN PERCEIVED BEHAVIORAL CONTROL (PBC)

Please click on the response that best reflects your perspective. (Response format was scaled on a 4-point Likert scale, ranging from 1: not at all true to 4: exactly true.)

1. I find it easy to be friendly with the environment.

2. I find it difficult to preserve resources and recycle. (R)

3. I am confident that I can protect the environment.

4. I can control my involvement in environmental preservation initiatives.

5. I am fully capable of protecting the environment.

6. Thanks to my resourcefulness, I always find a way to be friendly with the environment.

7. I am in full control of my actions to protect the environment.

8. I am good at leading a green lifestyle.

9. It is not easy for me to stick to my sustainability goals and preserve the environment. (R)

10. Being friendly with the environment is out of my hands. (R)

Differences in belief-consistent and belief-inconsistent learning in traditional college students

Carol Y. Yoder, Ruben Mancha, and Patrick Smith

Trinity University

ABSTRACT

Beliefs, described as adaptive mechanisms that frame experiences and shield against problems or criticism, impact learning and behavior. With maturation, adolescents and emergent adults are increasingly able to learn information inconsistent with their perspective, analytically and with deliberation. We hypothesized that upper-division traditional college-aged students should be more effective learning belief-inconsistent information relative to first-year college students. In three studies comparing first-year and upper-class traditional college aged students, participants read information about political issues, rated their opinion, and answered questions about issues. Results indicated that older students learned information contrary to their perspective better than consistent information, whereas two studies showed that first-years demonstrated better learning of information consistent with their beliefs. This suggests older students have better ability to control analytical reasoning. Over the span of only a few years, young adults provided age-related behavioral evidence of more complex comprehension and thinking. Our data suggests that experience and/or maturation can decrease the restrictive filter beliefs may have on learning.

KEYWORDS: learning, beliefs, adolescence, emergent adults, political reasoning

THE SCIENCE OF BEHAVIOR analysis often addresses socially important difficulties in children, adolescents and adults' behaviors, studying factors that reliably influence actions and reasoning. Much emphasis is placed on trying to solve problems that create significant issues for living. Not all change involves managing problems, especially in the first two decades where children and adolescents are continuing to grow and develop. In trying to understand challenges of late adolescence and emerging adulthood, one common goal of education is to achieve complex, higher order thinking. Indeed even by middle childhood, inquiry and argument are essential to developing effective habits of mind, related to life-long learning (Kuhn, 2008). With maturation, adolescents show increased ability to process complex ideas (Kuhn, 2006; Piaget, 1957). Decades of research have consistently shown that adolescence brings marked improvements in basic cognitive processes and higher order reasoning, as well as increases in speed, capacity, and expertise (Kuhn, 2006; Steinberg, 2005). Improvements in thinking are more evident with factual content (Blanchard-Fields, 1986) as emotionally involving information is more difficult to integrate. Better behavioral evidence of thinking also occurs when information runs contrary to belief (Klaczynski &

Narisimham, 1998). These developmental changes in learning and thought make understanding choices and behaviors more complex.

Over the last two decades imaging research has identified substantial changes occurring in the prefrontal cortical area during the time frame traditional college-age students begin university. Structural changes to brain provide a location for a range of behavior change and analysis (Schlinger, 2002). Between 17 and the early twenties (Giedd, et al., 1995; van der Molen & Molenaar, 1994) neural growth, pruning, and enhanced myelination all occur. Studies consistently find a non-linear reduction in gray matter coupled with a simultaneous increase in white matter in late adolescence and young adulthood (Blakemore & Choudbury, 2006). The decrease in gray matter is generally thought to reflect pruning of early adolescent synaptic proliferation—to remove unnecessary and unused synapses—refining neural communication. The white matter increase is presumed to improve efficiency, integration and regulation of cognitive and affective information. Additionally, while advancing into adulthood, individuals increasingly learn to control impulsive thoughts and drives (van den Bos, Westenberg, van Dijk & Crone, 2010). Indeed, many imaging studies have identified changes affecting areas in the brain regulating behavior and emotions and how we perceive and evaluate information (Paus et al., 1999). Blakemore and Choudhury (2006) speculated that adolescence ushers in a period of brain reorganization, where there is heightened sensitivity to cognitive and socially influenced experience and behavior.

Author note: Carol Y. Yoder, Department of Psychology, Trinity University; Ruben Mancha, Department of Finance and Decision Sciences, Trinity University, and Patrick Smith, Department of Psychology, North Texas State University. Correspondence concerning this article should be addressed to Carol Y. Yoder, Department of Psychology, Trinity University, San Antonio, Texas 78212, cyoder@trinity.edu, 512-999-8385, 01 512-999-8386 fax.

These structural changes set the stage for new possibilities particularly for emotionally engaging reasoning and behavior. Kunda (1990) explored how belief-motivated reasoning involves thinking about information consistent with one's beliefs, in ways that maintain and support those pre-existing beliefs. People often accept information consistent with their views while rejecting information contrary to perspective. Klaczynski and Narasimham (1998) found early adolescents dismissed evidence contrary to their religious views with analytic arguments. However, participants readily embraced supportive ideology with succinct and uncritical justifications. That is, ideas inconsistent with one's ideology were subjected to more rigorous scrutiny with more effortful deliberations. Similar patterns were found in adolescents' theories of homosexuality with theory-consistent and theory-inconsistent data (Kardash & Scholes, 1996). In a study comparing theories about beliefs and motivated reasoning biases, middle adolescents consistently demonstrated better reasoning than early adolescents, although biased thinking was similar (Klaczynski, 2000). This research underscores the necessity of considering beliefs, inherent theories and assumptions when trying to understand reasoning, at least for a substantial portion of Klaczynki's research sample. On the other hand, Klaczynski found that approximately one-third of his sample were more knowledge-driven and willing to push their personal theories to the side to consider evidence contrary to their perspective.

While there is work investigating early adolescence, it is difficult to find research exploring change in late adolescence (Steinberg, 2005). Further, in spite of the interesting possibilities suggested by imaging research, few studies have explored cognitive or behavioral concomitants. Thus, although there is solid evidence of structural change, changes in cognitive functioning and behavior have not been a focus of extensive investigation. Klaczynski and Cottrell (2004) suggested these structural changes should enhance one's ability to allocate resources. In discussing dual processes, they posited that this new structural foundation should enhance ability to intentionally engage effortful cognition, overriding automatic processes. Entertaining information consistent with one's ideas is likely to elicit automatic processing whereas considering information contrary to one's views normally promotes more analytic processing. Using analytic reasoning typically leads to rejecting evidence, especially with theory-inconsistent information due to cognitive dissonance mismatch. Essentially, Klaczynski and Cottrell suggested late adolescents have more control over whether or not they fully engage in critical thinking with the advent of this structural enhancement. While earlier development might have encouraged more focus on information that is readily available and consistent with one's knowledge, this increased biological substrate putatively provides the advantage of being able to more effectively engage information contrary to one's perspective, which requires more attention and effort. Taking a related tack, Kuhn (2006) noted the importance of increased intentionality and purposefulness as people develop better coordination of knowledge bases and more focused motivation to seek information and advance goals.

We compared learning performances on first-year and upper-class traditional-aged college students. We operationalized more complex learning by comparing belief-consistent and be-lief-inconsistent information, with the assumption that the latter requires more cognitive control and response inhibition. We were particularly interested in determining if there were differences in learning for belief-consistent and belief-inconsistent information, reasoning that, given the documented brain changes and experiences of college, advanced undergraduates should be more effective at dealing with more complex information (i.e., belief-inconsistent) relative to beginning undergraduates. To increase difficulty and tap into emotional regulation, we selected politically charged issues, of which most people have opinions. Because socio-political content is affectively charged for many people, it may trigger more automaticity and require more controlled processes and perhaps inhibition to override the initial reactivity (Klaczynski, 2004). Indeed Blanchard-Fields (1986) demonstrated that teens, young adults and middle-aged adults performed similarly when consolidating competing perspectives on neutrally-valenced issues. However, with issues that had more personal salience, such as an unplanned pregnancy, she found older (middle-aged) participants were more effective at balancing different points of view relative to young adults, who in turn were more effective than adolescents. Having the necessary cognitive resources to inhibit reactivity, along with a motivated interest in accuracy, should influence whether participants respond automatically or with greater intentionality.

Given the developmental change brought through experience and maturation in emergent adulthood, we might expect that younger college-aged participants, in comparison to older participants, would be more disadvantaged in a task that requires them to use controlled processes. Younger participants may have more difficulty selectively attending to relevant information, in part because they fail to inhibit their preferred response. They may also be less effective at learning and integrating content, especially when it is inconsistent with their beliefs. Here, beliefs reflect one's interpretation and response to environmental experiences and can be defined as behavioral phenomena that serve an adaptive function in helping us manage information or shield us from problems or criticism (Taylor & Brown, 1988). Beliefs also influence what information we attend to. When information differs from these interpretive frameworks, younger participants may be less effective in shifting from automatic to more controlled processing. Even in a traditional college sample, older college aged students may demonstrate more proficient comprehension of belief-inconsistent information from politically oriented vignettes, relative to younger college aged students.

Recognizing that if there were any differences between these closely related age groups they would be small, and variability between individuals would be great. We were concerned that individual differences would obscure subtle differences between groups. Individuals have different tendencies in how they structure their experiences to be meaningful; specifically, people vary in their disposition to evaluate and think (Cacioppo, Petty & Kao, 1984). Because willingness to comprehend politically oriented material might be affected by motivation for learning, we assessed interest in information seeking. We also were concerned about individual willingness to explore alternative perspectives. Closed mindedness may also impact how much effort would be expended on comprehending perspectives different from one's

own (Rokeach, 1954). As a consequence we also measured dogmatism, recognizing that political ideology is sometimes related to complexity in thinking (Tetlock, 1984).

» STUDY 1

Method

Participants. Seventy-four undergraduate students were recruited to participate from an introductory psychology class for course credit at a selective liberal arts college. Participants ranged from 17 to 24 years of age. A median split was used to divide our sample into two age groups, with the average age of our 43 younger participants being 18.64 (SD = 0.31) and 31 older participants' mean age was 20.77 (SD = 1.03, n = 31).

Materials

Dogmatism. Dogmatism has been described as a relatively closed system of beliefs about reality and absolute authority, which provides a framework for tolerance and intolerance of others (Rokeach, 1954). Behavior often labeled dogmatic is characterized by rigidity and inflexible methods of handling information, events, and people. We used the Trodahl and Powell (1965) 20-item dogmatism measure, where respondents rated how accurately each statement reflected their opinion on a 7-point Likert scale. A sample item representative of these items was "Most people just don't know what's good for them." Lower scores indicate more dogmatic styles.

Need for cognition. Need for cognition has been described as one's relative interest in engaging in thinking relative to an intuitive or experiential approach (Cacioppo & Petty, 1982). Behavior labeled as high need for cognition is characterized by more searching and attention to details and distinctions in categories. Cacioppo, Petty and Kao (1984) revised the original measure to 18-items, including statements such as "Thinking is not my idea of fun." Higher scores indicate more need for cognition.

Issue ratings. A brief description of a political issue was presented. After reading the description, participants answered three questions with Likert scales. The first question was (1) how much they agreed or disagreed with the issue (1 = high disagreement; 7 = high agreement), which we labeled 'belief.' The remaining questions asked (2) how important the issue was to them (0 = not important; 3 = very important), and (3) how much prior knowledge they felt they had about the issue (0 = low prior knowledge; 3 = high prior knowledge).

Eight political issues. Eight 250-word positions were written describing topical political issues. These vignettes covered an array of topics receiving media coverage in 2008, including presidential powers, health care reform, and immigration. Content included information and data that was not well known, and presumed to not already be known by participants. Half of the vignettes were written from a traditionally liberal viewpoint, half were written from a traditionally conservative viewpoint based on concepts discussed in newspapers and publications at that time (e.g., *Wall Street Journal, New Republic*). A variety of professionals representing both ends of the political spectrum provided feedback and suggestions to maximize the appearance of reasonability and even-handedness regarding the persuasive content of the vignettes.

Assessment questions. Five free response questions were used to assess information described in each vignette as a measure of comprehension. It was unlikely that participants could answer these questions without reading the vignettes.

Procedure

After providing informed consent, participants completed two pre-test assessment measures, a need for cognition scale and a dogmatism scale. Participants sat in front of a computer, presented with descriptions of eight political issues. Participants were asked to rate each issue on three descriptive questions. Participants were then presented with a series of eight vignettes, presented in a randomized and counterbalanced order. Comprehension was assessed after all vignettes were presented.

Learning assessment. After participants had read all vignettes, a series of follow-up free response questions were displayed. For example, after reading the vignette focused on presidential powers, participants typed in responses to questions such as "Besides the Constitution, what other documentation by the Founding Fathers supports the executive branch's power? How does this policy provide flexibility for governance?"

Data coding. Raters coded individual responses to follow-up questions on a 0 to 4 point scale, with higher numbers representing better quality, more inclusive answers, that incorporated material presented in each vignette. The first author created a scoring codebook with conceptual examples of five levels of performance and detailed descriptions of representative answers particular to each vignette. A total of five undergraduates were trained and two raters were assigned to score items from each of the vignettes. Coding responses were compared with κ. Rater reliability was good, with an inter-rater agreement of 0.81 (Landis & Koch, 1977); disagreements were resolved through discussion between raters. Scores from all items assessing a particular vignette were summed, divided by the number of items and converted to a percentage.

We were interested in comparing comprehension of issues most strongly consistent and most strongly inconsistent with each participant's beliefs. As it would be expected, participants found different issues more or less compelling. Participants' initial rating of their perspective was used as the criteria. That is, the most extreme rated vignettes (strongest agreement, strongest disagreement) represented the most positive and negative issues for that individual and their respective learning score on that particular vignette was entered as a dependent variable. All extreme issues were also rated as having importance. Rated importance was also considered, but as a check to be certain the respondent viewed the issue as salient. After determining the appropriate issues for each participant (based on their rating), a *learning assessment score* (LA) was created which represented their percent correct score for that issue. This involved summing up scores for each item assessing that vignette, so that a maximum score of 20 points was possible (5 questions × 4 points), which was then converted to percent correct. Overall percent correct was initially used to look at comprehension of the eight issues. When agreement and importance ratings were identical on more than one issue, the vignettes' performances were averaged to create the LA, which was used as the dependent variable.

Table 1. mean ratings of political commentary descriptors

issue	correct % M	correct % SD	rating M	rating SD	importance M	importance SD	knowledgeability	time
estate taxation*	67.67	18.31	5.60	1.39	1.29	0.75	0.57	93.871
assisted suicide*	58.94	15.89	4.02	2.12	2.03	0.67	1.22	65.186
stem cell*	52.76	19.58	3.48	2.08	2.31	0.78	1.31	102.36
environmental protection	56.90	16.99	5.14	1.59	2.36	0.81	1.26	80.734
presidential powers*	53.13	19.70	3.52	1.82	1.86	0.81	1.17	83.240
health reform	32.87	18.86	5.38	1.12	2.26	0.72	0.91	113.85
immigration	47.24	16.17	5.19	1.59	2.05	0.66	1.16	91.464
affirmative action	58.30	15.15	4.50	1.74	1.79	0.77	1.34	81.437

Note: Personal ratings ranged on a Likert scale from 1 (strongly disagree) to 7 (strongly agree). Personal ratings of issue importance ranged from 0 (not important) to 3 (very important). Ratings of knowledgeability on each issue ranged from 0 (not informed) to 3 (very informed). Time spent on each issue was recorded in seconds. *Vignette content conveyed a more traditional, conservative perspective. The non-asterisked vignettes were framed from a progressive, liberal stance.

Two learning assessment scores were created: *LA-A (agree) and LA-D (disagree)*. The former represented learning about 1) the issue with which a given participant reported was most consistent with belief (most strongly agreed issue, LA-A) and 2) a second issue which was most inconsistent with belief (most strongly disagreed issue, LA-D). Dependent variables (LA-A, LA-D) were distinct for participants, and were based on their stated position on vignette issues they most strongly agreed, or disagreed with. As described in the previous paragraph, a maximum score of 20 was converted to a LA-A and LA-D percentage. Because of individual vignette variations in quality of responses to follow-up questions, performance scores were converted to *z*-scores. This method allowed comparison of performance on issues most consistent and most inconsistent with their personal beliefs.

Results

There was considerable diversity of opinion on the issues we targeted. Estate tax, assisted suicide, presidential powers, and stem cell research funding were framed from a conservative point of view whereas affirmative action, environmental protection, health reform and immigration reform were framed from a liberal point of view. There was a very limited range of ratings about pre-existing knowledge, with participants being least knowledgeable about estate tax laws ($M = 0.49$; $SD = 0.55$) and most knowledgeable about stem cell technology ($M = 1.3$; $SD = 0.61$), with an average informed rate of 1.05 ($SD = 0.42$). Health reform was the most poorly understood issue although more time was spent trying to comprehend its content than any other issue. The taxation issue was best understood although participants studied it an average amount of time. See Table 1 for additional information.

To assess whether type of content was handled differently with increased age, we used a mixed ANOVA design with age (first-years, seniors) as a between subjects factor and belief (LA-A; LA-D) as a within subjects factor. We found a main effect for age, $F(1, 64) = 6.96$,

$MSE = 0.9498$, $p = 0.011$, $\eta_p^2 = 0.09$, with older students performing better than younger students. We also found an interaction between age × belief, $F(1, 64) = 6.00$, $MSE = 3.26$, $p = 0.0171$, $\eta_p^2 = 0.09$, such that younger participants substantially improved when challenged with content inconsistent with their beliefs. Nevertheless, older participants were more effective with both belief-consistent and especially belief-inconsistent content. Comparing time spent on task, the groups were similar, $F(1, 73) = 1.48$, $p = 0.23$.

There were differences between older and younger participants in terms of need for cognition, $t(68.94) = 2.43$, $p = 0.01$, $d = 0.55$, with older respondents indicating a higher need for cognition ($M = 90.31$, $SD = 9.67$) than younger respondents ($M = 83.58$, $SD = 13.67$). Partitioning need for cognition into three levels (high, medium, low) also yielded differences in overall learning performance, with individuals with higher need for cognition showing better comprehension, $F(2, 64) = 2.97$, $MSE = 0.366$, $p = 0.05$, $\eta_p^2 = 0.09$. Level of dogmatism also differed between groups, $t(66.712) = 2.18$, $p = 0.03$, $d = 0.49$, with younger participants averaging lower scores ($M = 74.21$, $SD = 10.92$) relative to older participants ($M = 78.92$, $SD = 7.00$), suggesting less flexibility in younger participants' values. Learning performance did not differ between median-split groups of high and low dogmatism. Both groups learned equally content consistent and inconsistent with their reported belief ($F(1, 65) = 0.22$, $p = 0.64$). See Table 2 for additional information.

Discussion

As suggested by Klaczynski (2004), our data show that older college students were better able to respond to questions about information that was contrary to their own point of view. Older respondents demonstrated better learning of contrary information, relative to information congruent with their pre-existing perspectives, and as compared to younger participants. Also, younger students benefited more from the incongruent information relative to information consistent with their beliefs. Perhaps as Klaczynski has repeatedly suggested, belief-incongruent content is generally more engaging at least in the case of motivated emergent adults who have chosen a liberal arts environment for study. While these results were promising, we were interested in developing a more strategic upper-class (junior/senior) sample, as opposed to simply studying people finishing their general education classes in introductory psychology. Also, we were concerned about the amount of time required to complete the survey and respondent fatigue, so the number of vignettes was reduced to six. (Selecting

Table 2. descriptive information about primary measures

group	age	total %	SD	belief-con	SD	belief-inc	SD	dogma	SD	needcog	SD
younger	18.6	49.26	13.66	41.96	24.24	49.19	23.84	74.21	10.92	83.57	13.67
older	20.8	55.46	13.92	49.70	25.73	54.72	24.86	78.92	7.02	90.31	9.670

which vignettes to exclude also reflected a change in government stem cell policy and lower perceived importance of tax reform.) Two additional questions were added to better assess learning of each vignette. Study 2 repeated the procedures used in Study 1, except it included fewer (six) vignettes, two additional questions per vignette, and we specifically recruited upper-class students and first-years for our sample. As before, vignettes were presented to participants in a randomized and counter-balanced order and all questions pertaining to a vignette were grouped together.

» STUDY 2

Method

Participants

Eighty-three college undergraduates volunteered to participate in a psychology study for course credit at the same liberal arts institution. Participants were drawn from undergraduate psychology courses and upper-division seminar courses. Although efforts were made to make each vignette plausible and reasonable, participants had to show variability in beliefs to provide data for analyses. Those who had little or no range in agreement ratings were omitted. That is, respondents had to have at least three different ratings across all six vignettes to meet criteria for inclusion so that a belief-consistent (positive) and belief-inconsistent (negative) issue could be contrasted (from a middle or neutral agreement rating). Additionally, because some subjects did not demonstrate the willingness to read, think and provide thoughtful responses critical in this study, we dropped participants whose overall performance was poor ($z < -1.5$). Removing unmotivated participants and those without opinions left 36 younger students ($M = 18.6$, $SD = 0.31$) and 17 older students ($M = 21.2$, $SD = 0.98$) in the sample. Since we were focused on learning, omitting those who chose not to engage in this demanding task was both reasonable and necessary to test our hypothesis.

Measures and procedure

After participants had read all the vignettes, a series of follow-up free response assessment questions were asked (seven questions per vignette). As before, agreement rating was used to determine which vignettes' scores were used to represent the independent variable of belief, with belief-consistent (agree) and belief-inconsistent (disagree) providing the two levels of comparison. If more than one issue was rated extremely, importance ratings discriminated which vignettes' performance was used. In the case of equal issue ratings (agreement and importance), percent correct was averaged across the appropriate vignettes.

Learning assessment

As before, using an updated and substantially revised scoring codebook, raters coded responses from 0 to 4, with higher numbers representing better quality, more inclusive answers. Rater

reliability was acceptable, $\kappa = 0.84$ and disagreements were resolved through rater discussion. Once the belief-consistent and belief-inconsistent issues were determined, that vignette's items were summed, divided by the number of items and converted to a percentage. These percentages, LA-A, and LA-D, were treated as dependent variables and converted into z-scores.

Results

To assess the relationship between belief, age and performance, we conducted a repeated measures ANOVA with age (2: first-years, upper-class students) as a between subjects factor and belief (2: LA-A; LA-D) as a within subjects factor. This revealed an belief × age interaction ($F_{(1, 40)} = 3.97$, $MSE = 3.08$, $p < 0.05$, $\eta_p^2 = 0.09$). First year students performed similarly on content that was consistent with their perspective whereas more advanced students' performance was relatively higher when answering questions based on vignettes with which they disagreed. Younger students performed similarly on both types of content but older students were more effective with information inconsistent with their perspective. Interestingly, this suggests that approximately two and one-half extra years of experience may help motivated performers direct their attention to content that requires more effort to comprehend. In comparing time spent on task, the older group spent significantly less time on task, $F_{(1, 40)} = 10.25$, $p < 0.05$, $\eta_p^2 = 0.19$ although no statistical difference was identified between the time spent on belief-consistent information as compared to the belief-inconsistent information.

Because of concerns about other individual differences between groups, dogmatism and need for cognition were also assessed and evaluated. The repeated measures ANOVA was evaluated with dogmatism and need for cognition as covariates. Neither of these measures differed in this sample of younger and older college participants, nor did these measures of individual differences affect learning performance. Table 3 presents additional descriptive information contrasting the two age groups.

Discussion

As suggested by other researchers' data and theorizing (Klaczinski, 2000; 2004; Klaczinski & Cottrell, 2004; Kuhn & Pease, 2006), first-year students were more proficient learning content that was consistent with their perspective and easier to cognitively digest. Content that was belief-inconsistent was handled more effectively by more advanced students. There was no main effect for age in this sample.

This study provides behavioral evidence of more effortful learning in traditional upper-class undergraduates relative to their first-year peers on belief-inconsistent information. Differences emerged such that slightly older undergraduates handled content contrary to their perspectives better than content congruent with their beliefs and much better than first-years. In contrast, first years learned most effectively information consistent with their personal beliefs. Other individual difference measures did not consistently predict performance, pointing to the critical nature of motivation and effort as Kuhn (2006) underscores.

Table 3. subject variables and belief-consistent and belief-inconsistent z-scores (*SD*)

group	belief-con	time	belief-inc	time	dogma	need cog
younger	0.033 (0.94)	101.6 (37.26)	0.02 (0.87)	94.57 (45.22)	75.54 (8.56)	71.43 (5.61)
older	0.140 (1.10)	80.86 (34.13)	0.26 (0.78)	75.51 (31.74)	78.20 (10.4)	71.19 (6.28)

Although impossible to separate experiential learning (which we believe college provides) from maturation, our findings provide evidence for cognitive advancements that fit with biological changes reported by others (Blakemore & Choudury, 2006; Giedd et al., 1995; Giedd et al., 1999; Paus, et al., 1999). Data from study 2 also found more complex thinking in a relatively homogeneous traditional college-aged sample comparing young and slightly older students.

Study 3 utilized similar materials and procedures. In this study we added a political party affiliation item to our demographic questions. Because both working memory and verbal comprehension could mediate relationships previously seen between age and belief-based learning, a working memory measure, and a verbal intelligence measure were added. As well, there may have been valence-specific processes which were engaged due to the emotional nature of the content in the vignettes (i.e., strong beliefs, both in agreement and disagreement). Neutrally-rated issues might engage differential learning processes or differential motivation to spend time and energy reading and answering questions about specific issues, so in study 3 we planned to compare the most neutral issue with those issues that elicited the strongest opinions. Dogmatism and need for cognition scales were omitted since they were ineffective in study 2.

Research indicates that conceptual content should be more accessible to more mature learners (Chapman, Gamino, & Mudar, 2012; Kuhn, 2006). We varied type of question (factual/conceptual) with at least 3 conceptual and 3 factual questions assessing learning about an issue. After determining which issues best characterized belief-consistent, belief-neutral, and belief-inconsistent perspectives for a given individual, we separated factual and conceptual items to assess whether conceptual content was better handled by older students.

» STUDY 3
Method
Participants
One-hundred and five college students participated in this study to complete course requirements, however eighteen participants were excluded from the analysis because they failed to complete the task satisfactorily using criteria from study 2 (z-score for total performance < −1.5). Younger students ($n = 42$; $M = 18.9$, $SD = 0.6$) were recruited from a student pool made up of introductory psychology students, and older students ($n = 45$; $M = 21.5$, $SD = 0.7$) were recruited from upper-division courses. In both instances participants received class credit. Forty-seven percent self-identified as Democrats, 14% as Republicans, 10% as Independents, 5% marked Other and 24% claimed no party affiliation.

Additional material
N-back test of working memory capacity. In the *n*-back test of working memory a test trial was included to familiarize participants with the procedure. Participants were presented with a series of words and were told that they would have to type when a word was presented "*n*" words prior to the current word. This task involved keeping a running list of words "in mind" in order to check current word presentation against prior word presentation.

For example, a participant completing a 2-back task would signal a response on the second presentation of the word "flower" in the four-word series block-*flower*-computer-flower. Following the trial test, participants encountered four series of 20 words presented in serial order. For the first two word sets participants were instructed to indicate when a word was presented "2-back", whereas participants were instructed to indicate when a word was presented "3-back" for the latter two word sets. To reduce the possibility that visual representations of the words would enhance *n*-back performance, words similar in meaning were selected using Thesaurus.com and all words were verbs. *N*-back task performance was assessed using three variables; (1) the number of correct responses, (2) errors of commission (false response), and (3) errors of omission (no response).

Quick test (verbal intelligence measure; Ammons & Ammons, 1962). This vocabulary measure involves reading a vocabulary word and clicking on one of four pictures which best represents the concept. After three consecutive misses, the measure ends; otherwise, there are 24 possible matching items.

Procedure
As in study 1 and 2, after providing consent, participants rated their belief towards a variety of political issues & rated their existing knowledge and interest. Participants then responded to the *n*-back working memory test. Next, participants read six political perspectives presented in randomized and counterbalanced order on a computer screen. After completing these measures, participants were asked factual and conceptual follow-up questions on each political perspective. After responding to the follow-up questions, the Quick Test verbal intelligence measure was presented.

As in the previous studies when measuring the follow-up questions, raters coded responses on a 0 to 4. Higher numbers represented better quality, more inclusive answers, that incorporated material presented in each vignette. In addition to comparing belief-consistent and belief-inconsistent performance, we also included performance on the issue rated most neutral and at least of some importance. Inter-item consistency within political perspective was reasonable ($\kappa = 0.81$), so items were separated into percent of factual content that was correct and percent of conceptual content that was correct. Percent correct scores were sorted by belief (agree, disagree, neutral) and converted into z-scores.

Results
To assess the relationship between belief and age, a repeated measures ANOVA with age (2: first-years, upper-class undergraduates) as a between subjects factor and belief (3: LA-A; LA-D; LA-N) and Content (2: factual, conceptual) as within subjects factors. Using the *n*-back commission score as a covariate, there was an interaction between belief × age, $F(2, 83) = 3.15$, $MSE = 2.10$. $p = 0.05$, $\eta_p^2 = 0.05$. First-years did relatively better on content that fit with their perspectives, whereas upper-class students did better on content contrary to their point of view. Contrary to Kuhn and Pease (2006), conceptual content was not differentially remembered in this content comparison nor were there other interactions. Only the commission working memory measure was an effective covariate in this analysis.

Table 4. descriptive statistics and *z*-score learning averages by group

group	belief-con	belief-neu	belief-inc	verbal IQ	*N*-back correct	*N*-back omiss	*N*-back commis
younger	−0.11 (0.92)	−0.27 (0.94)	−0.18 (1.01)	15.31 (2.60)	18.68 (2.73)	7.44 (2.65)	2.31 (2.31)
older	0.33 (0.98)	0.01 (1.07)	0.41 (0.87)	17.51 (2.68)	16.82 (2.41)	8.26 (2.36)	2.00 (2.05)

In a secondary analysis of age groups, *t*-tests were used to compare age groups. Older students scored higher on verbal intelligence, $t(85) = -2.40$, $p = 0.02$ while younger students scored higher on *n*-back hits, $t(85) = 2.69$, $p = 0.01$. *N*-back omission rates also differed, $t(85) = -2.56$, $p = 0.01$ with older students having more errors. Errors of commission were similar for both groups.

Discussion

Similar to study 2, study 3 also found younger students did more poorly on content that was contrary to their opinion relative to belief-consistent information, whereas older students consistently did better on content that was contrary to their perspective. Neutral content was handled relatively similarly to agreed-with content for first years, but was remembered least well by older students. This difference in handling neutral information suggests interest in any particular issue is central to performance, especially for upper-class students.

Working memory showed little relationship to political issue learning. Only one of three *n*-back measures was predictive of learning performance. It is not clear why commission mattered but it could be interpreted as an indication of whether or not working memory was fully deployed to the task at hand. In general there are several indicators that the older students were less engaged in the task (e.g., incomplete responses to questions, *n*-back hits, *n*-back omissions). On the other hand, perhaps this would be expected given the relatively homogeneous group of participants, particularly when poor performers were removed from the sample.

While we expected that factual content would be easier to remember and would be more easily recalled by both age groups, our operationalization of factual content did not yield age differences. As well, it was expected that conceptual content would be better processed by older students, and this too, did not appear to exert significant influence on learning ability. The failure to find a difference in performance was likely due to insufficient distinctions in these divisions with our questions. Contrary content once again inspired better performance from slightly older students even though they did not demonstrate better working memory relative to their marginally younger counterparts.

Although it is not possible to separate experiential learning from maturation, our findings are suggestive that at least experiential advances are associated with more complex thinking in a relatively homogeneous traditional college-aged sample. As developmental scholars increasingly embrace the notion of emergent adulthood, these changing potentials highlight the distinctiveness of this time frame and the importance of systematic developmental study (Arnett, 2000).

» GENERAL DISCUSSION AND SUMMARY

In three studies exploring learning about emotionally-charged political information in traditional college-aged students, we found that first-years and upper-class undergraduates differed in the kind of content they learned best. Our inquiry was inspired by the structural changes in brain identified in young adults that should result in improved ability to self-regulate, monitor, and decide whether or not to allocate additional resources to comprehending effortful content. We created tasks that we hoped would correspond to skills that emanate putatively from these structural changes. As suggested by developmental researchers using imaging technologies, we found that slightly older traditionally-age college students had the potential to be more effective at dealing with political information contrary to their opinion, relative to younger students.

While using college students as a prototype for adult cognition and behavior is generally recognized as problematic, substantial developmental milestones occur during the time frame of college years. Imagery-identified brain changes during these periods make it even more important to reconsider sampling issues based on age (i.e. first year students vs. more advanced students), in addition to generalizability to adult samples. It is more than a bit ironic that college students have provided samples for many advances in psychological knowledge and yet it is difficult to find any behavioral or cognitive data comparing development or maturational changes during this time frame. There are several reasons for this. Of course, one reason is that no one looked—for many decades the assumption was that the brain was more or less fully developed in adolescence with only informational and experiential updates. The other issue is that even if behavioral and development changes were scrutinized, there is so much variation within a group of individuals that it is difficult to sift through. These issues are further amplified by participant interest in optimizing performance, especially when tasks are difficult and require substantial cognitive resources. Additionally, performances on difficult tasks are especially affected by many other situational variables. Indeed, one might reasonably expect to find small but predictable developmental differences only on more challenging tasks.

In a younger adolescent sample, Klaczynski (2000) found three common patterns in the paradigm he used to study how pre-existing beliefs might affect reasoning in younger adolescents. Some adolescents were biased towards groups to which they were affiliated, others were biased towards groups to which they were *not* affiliated, and a third group's interest in knowledge superseded commitment to a particular belief. Except for this knowledge-based group, Klaczynski (2000) stated that identifying patterns in reasoning was highly dependent on knowing adolescents' assumptions and beliefs. Consistent with Kuhn's and

Klaczynski's theorizing, our data suggest that maturation might increase the likelihood of people choosing to consider evidence from an analytical perspective, especially when it runs contrary to their existing perspective. While individual differences might seem important, in terms of preferences, approaches, values, intelligence (Stanovich & West, 1997) and scientific reasoning competence (Klaczynski, 2000), these factors did not appear to be a consistent predictor of learning outcomes. Additionally, political issues are not of interest to everyone, and this is particularly true for many emergent adults. Longitudinal data or more homogenous samples may provide a better assessment of how belief-consistent and inconsistent information is handled.

Consistent with Kuhn's and Klaczynski's theorizing, our data suggest that maturation might increase the likelihood of people choosing to analytically consider evidence, especially when it is contrary to their existing perspective. While individual differences might seem important, in terms of preferences, approaches, values, intelligence (Stanovich & West, 1997) and scientific reasoning competence (Klaczynski, 2000), these factors did not provide a consistent predictor of learning outcomes. Additionally, political issues are not of interest to everyone, and this is particularly true for many emergent adults. Longitudinal data or more homogenous samples should provide a better assessment of how belief-consistent and inconsistent information is handled.

One limitation of this study is that it requires participants to work. The effect does not consistently occur without attention to the task. We recognize there may be other salient parameters as well, but motivation is key, as Kuhn (2006) notes. Second, the results could be driven more by experiential factors associated with advancement though a rigorous college curriculum than by neurodevelopment. It seems likely that the experience of attending college enhances one's ability to deal with contradictory information simply through a practice effect. To further explore this possibility, it would be necessary to include a sample of similarly aged participants who had little post-secondary education but were also highly motivated to learn.

It is also useful to consider age-independent effects (i.e. practice effect) that might be associated with attending college. There likely exists an interaction between age and the experiential factors associated with the college curriculum. Akin to visual development, there could be certain specific inputs that are required while the brain is in a "critical" developmental period. Given that substantial plasticity exists in the prefrontal cortex of developing young adults, the college experience could provide the necessary input that might foster further development of prefrontal circuits involved in emotion inhibition and learning. The cumulative action of all of these factors might culminate in older college students performing better on tasks that require them to attend to information contrary to their own opinion.

We believe this information could be utilized by those in the business of training, persuasion, or education. For instructors designing college curricula, it could be utilized to shape content presentation. While the lower- and upper-division course distinction already informally recognizes these developmental learning changes in young adults, an argument can be made for introducing belief-inconsistent information earlier. The way information is framed could theoretically make those in lower-division classes more effective at learning belief-inconsistent information and more engaged in taking apart weak arguments or poorly constructed evidence. Even though belief-inconsistent information is not as well attended to by these younger students, framing belief-inconsistent information in a way that is more congruent with existing beliefs could facilitate learning and critical thinking skills.

The implications for how adolescents handle information contrary to belief, are suggestive. Younger people are more likely to disregard belief-inconsistent information, instead selectively attending to information that is consistent with their existing beliefs. Individual differences in working memory, verbal intelligence, need for cognition and dogmatism did not explain these learning performances. Unfortunately, disregarding contrary information can lead to errors in reasoning and decision-making, resulting in less-than optimum behaviors. Emerging adults, on the other hand, are better equipped to navigate contrary opinions and information, thus making better choices possible.

Additional attention should be focused on behavior and thinking during this critical age where young adults are emerging as new, and increasingly influential members of society. Cross-sectional comparisons here yielded small differences whereas longitudinal data, especially when paired with imaging research, would be more illuminating. Decision making involves two different yet related cognitive processes—information search and information integration. After information is perceived and learned, it must be interpreted and either integrated into an existing belief, or discounted. In research on intelligence tests, Frey (1981) found that some participants engaged in compensatory behavior—changing self-judgments due to belief-incongruent information—as an alternative to confirmation bias. While Frey (1981) explained this phenomenon as a way to avoid cognitive dissonance, it can also be seen as integrating conflicting information into a new or existing belief.

Our research focused on navigating and learning belief-consistent and belief-inconsistent information in late adolescence and emergent adulthood. While increased maturation does not necessarily mean more effortful processing will be chosen, experience and identified structural change (e.g., Giedd et al., 1995; 1999) makes the opportunity for critical thinking more available. Given individual differences, some individuals will choose to bypass thinking and respond automatically, but others will decide to navigate the flow of contrary information, thereby increasing the likelihood of an optimized choice and behavior. ■

REFERENCES

Ammons, R. B., & Ammons, C. H. (1962). The Quick Test (QT): Provisional manual. *Psychological Reports*, *11*(1), 111-161.

Arnett, J.J. (2000). Emerging adulthood. A theory of development from the late teens through the twenties. *American Psychologist, 55*, 469-480.

Blakemore, S-J & Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry, 47,* 296-312.

Blanchard-Fields, F. (1986). Reasoning on social dilemmas varying in emotional saliency: An adult developmental perspective. *Psychology & Aging, 1,* 325-333.

Cacioppo, J.T., & Petty, R.E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42,* 116 − 131.

Cacioppo, J.T., Petty, R.E., & Kao, C.F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48,* 306 − 307.

Chapman, S.B., Gamino, J.F. & Mudar, R.A. (2012). Higher order strategic gist reasoning in adolescence. In V. Reyna, S. Chapman, M. Dougherty & J. Confrey (Eds.) *The Adolescent Brain: Learning, Reasoning, and Decision Making.* (pp. 123-151).Washington, DC: APA.

Frey, D. (1981). The effect of negative feedback about oneself and cost of information on preferences for information about the source of this feedback. *Journal of Experimental Social Psychology, 17(1),* 42-50.

Giedd, J. et al. (1995). Reliability of cerebral measures in repeated measures in repeated examinations with magnetic resonance imagery. *Psychiatric Research: Neuroimaging, 61,* 113-119.

Giedd, J. et al. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience, 2,* 861-863.

Kardash, C.M. & Scholes, R.J. (1996). Effects of pre-existing beliefs, epistemological beliefs, and need for cognition on interpretation of controversial issues. *Journal of Educational Psychology, 88,* 260-267.

Klaczynski, P.A. (2000). Motivated scientific reasoning biases, epistemological beliefs, and theory polarization: A Two-process approach to adolescent cognition. *Child Development, 7,* 1347-1366.

Klaczynski, P. (2004). A dual-process model of adolescent development: Implications for decision making, reasoning, and identity. *Advances in Child Development and Behavior, 32,* 73-123.

Klaczynski, P.A. & Cottrell, J.M. (2004). A dual-process approach to cognitive development: The Case of children's understanding of sunk cost decisions. *Thinking & Reasoning, 10*, 147-174.

Klaczynski, P.A. & Narasimham, G. (1998). The development of self-serving reasoning biases: Ego-protective versus cognitive explanations. *Developmental Psychology, 34*, 175-187.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*, 480-498.

Kuhn, D. (2006). Do cognitive changes accompany developments in the adolescent brain? *Perspectives on Psychological Science, 1*, 59-67.

Kuhn, D. (2008). *Education for Thinking.* Cambridge: Harvard University Press.

Kuhn, D. & Pease, M. (2006). Do children and adults learn differently? *Journal of Cognition and Development, 7,* 279-293.

Landis, J.R. & Koch, G.G. (1977). The measurement of observer agreement for categoritcal data. *Biometrics, 33,* 159-174.

Paus, T. et al. (1999). Structural maturation of neural pathways in children and adolescents in vivo study. *Science, 293,* 1908-1911.

Piaget, J. (1957). *Construction of reality in the child.* London: Routledge & Kegan Paul.

Rokeach, M. (1954). The nature and meaning of dogmatism. *Psychological Review, 61*, 194-204.

Schlinger, H.D. (2002). Concepts in behavioral development. *Behavioral Development Bulletin, 1*, 1-8.

Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends in Cognitive Sciences*, 9(2), 69-74.

Taylor, S.E. & Brown, J.D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103* (2)*,* 193-210.

Tetlock, P.E. (1984). Cognitive style and political belief systems in the British House of Commons. *Journal of Personality and Social Psychology*, *46*, 365-375.

Trodahl, V.C. & Powell, F.A. (1965). A short-form dogmatism scale for use in field studies. *Social Forces, 44,* 211-214.

van den Bos, W., Westenberg, M., van Dijk, E., & Crone, E.A. (2010). Development of trust and reciprocity in adolescence. *Cognitive Development*, 25(1), 90-102.

van der Molen, M.W. & Molenaar, P.C.M. (1994). *Human behavior and the developing brain*. In G. Dawson & K. Fischer (Eds.) (pp. 456-490). New York: Guilford Press.

The desires that were denied: (Re-) construction of sexual identity in middle adulthood

Thomas B. Swan[1] and Suzie Benack[2]

1 Siena College
2 Union College

ABSTRACT

Many men who entered adulthood in the 1970's through 2000 experienced sexual desire for other men in their adolescence, but did not integrate this experience into their identities. After forming heterosexual identities and entering heterosexual marriages, some experience a re-emergence of same-sex desire in midlife. We examine posts to three online groups for such men in "mixed-orientation marriages" to describe the ways in which the inadequacies of available cultural scripts for sexual orientation impede their ability to re-integrate their same-sex desire into their adult identities. We also suggest that the men who can make use of advanced forms of adult cognition are better equipped to transcend the limitations of cultural scripts and form a more coherent and inclusive adult identity.

KEYWORDS: gay married, mixed orientation marriage, midlife, adult identity, dialectical thinking

TRADITIONAL MODELS OF IDENTITY formation assume that in adolescence, ideally, a relatively coherent and cohesive identity can be created. This provides a stable platform for adult development, bringing important aspects of the self into a relatively harmonious relationship and connecting the individual to her social context in a meaningful and continuous way (e.g., Erikson, 1968; Marcia, 1994; Westen, 1985). This developmental task most often cannot be achieved with both complete scope and complete coherence. However, it is unlikely that adolescent identity will both include all potentially relevant aspects of the self *and* to create an identity that is unified and which harmoniously relates all the parts. It is necessary, then, that some aspects of the self that cannot be easily integrated with others be "sacrificed" (consciously or unconsciously) in order to form a more coherent identity. This compromise is necessary to make initial young adult life choices possible, but it also means that those choices will be, to some extent, ambivalent and conflicted.

Some of the forces that create challenges in integrating aspects of the self into a coherent identity include

» intrapsychic conflicts—aspects of the self which are incompatible, e.g. a desire for grand achievement vs. a desire to avoid risk of failure;

» interpersonal conflicts—identity-salient others who will be hurt, disapproving or angry if the adolescent acts on certain motives;

» social forces which make motives difficult to enact or impose costs for acting on them—limitations due to one's economic resources or social status, social taboos sanctioning particular behaviors in general or for members of certain social groups (e.g., if a person has aspirations which are economically unreachable or deemed "inappropriate" for someone of their social status);

» cognitive structures which limit the ability to integrate seemingly discrepant features of a self-system—e.g., a mode of understanding in which it "doesn't make sense" to be both a "good" daughter and angry at one's mother.

For many people, adolescent formulation of identity and early adult life choices form a more or less adequate basis for later adult identity and life, with gradual adaptation and elaboration in response to new life circumstances and personal growth. In some cases, however, the motives which were sacrificed/excluded from identity re-emerge later in life with a suddenness and intensity that creates a sense of crisis. Recognizing re-emerging parts of the self can create intrapsychic conflicts (*What I'm feeling conflicts with who I am),* conflicts in one's life structure (*I made choices/got where I am based on not feeling this; these feelings don't fit in my current*

life), and conflicts with one's relational commitments (*I created bonds and made promises based on not feeling this*).

Several "expectable events" in the course of adult life can alter the balance of psychic forces present in adolescence in ways that contribute to re-emergence of previously "sacrificed" motives. Below is a list of examples.

» Changes in the internal hierarchy of motives—Motives which were dominant in adolescence and early adulthood become less pressing, in part because their goals have been achieved.

» Changes in the interpersonal context of choice—Important relationships that led force to sacrificing a motive become less relevant as the relationships are redefined; in particular, relationships to parents can become less salient forces as one becomes more independent and parents age and/or die.

» Changes in social norms—Especially in post-modern cultures with rapid rates of social change, the social forces that shaped identity formation in adolescence may change dramatically during adulthood. Motives whose expression was prohibited become socially acceptable, and a wider range of actions become possible without the social costs they would have incurred at an earlier point in one's life.

» Changes in cognitive structure—The development of more complex cognitive structures in adulthood can render the "unthinkable" (in adolescence") something that now "makes sense," and thus becomes possible.

» Recognition of potential finality of the sacrifice as mortality becomes more salient—As an adult increasingly experiences his lifespan as finite, and his remaining time as limited, it can feel more urgent to recover what has been "missing" in his earlier life.

Any of these forces can lead to a crisis of identity in adulthood (a "midlife crisis"), in which the re-emergence of previously denied desires lead to re-opening fundamental identity questions and a re-evaluation of the life plan. Confronting questions of identity as an adult in contemporary American culture can be frightening and demoralizing, for several reasons. The dominant lifespan narrative in American culture includes and supports identity crisis in adolescence, but expects adults to "have it together" enough to make stable commitments to work and relationships. Being unable as an adult to say what one wants or to tell others what they can count on, then, can be experienced as a personal failure or a sign of immaturity. Moreover, American culture does not provide adults with the social supports given to adolescents in identity crises. A period of moratorium in which to experiment with temporary roles, values, and relationships, (Erikson, 1968), social approval for rejecting others' expectations and values in order to achieve personal authenticity (Erikson, 1968; Marcia, 1994), and tolerance of temporary self-centeredness, self-absorption, and irresponsibility. People renegotiating identity in adulthood must resolve identity issues while embedded in adult relationships of economic and social interdependence and responsibility. Their spouses, children, family, friends, and colleagues share the costs of the process and have a stake in its outcome. The demands of adult life limit the time and degrees of freedom available for exploration and experimentation. They also pose moral issues. While contemporary Western cultures encourage freedom, tentativeness about commitment, and self-focus in adolescents, adults who exhibit these traits are often criticized as being selfish, immature, irresponsible, or even pathological.

In this paper, we examine a particular group of midlife adults who experience the re-emergence of previously denied motives. That is men who experienced erotic attraction to other men in adolescence, but did not form "gay identities." These men entered heterosexual marriages, but found themselves, in middle adulthood, once again feeling intense same-sex desire, to the degree that it propels them to reconsider their sexual identity and their life structure. These men are often called "gay/bi-married men." For reasons which will become clear later in the paper, we want to avoid characterizing their sexual desires in this way, so will refer to them as SSA spouses—spouses with a same-sex attraction. In addition to giving a general description of the adult identity crisis precipitated by the re-emergence of same-sex desire, we will focus on the ways in which the inadequacy of dominant social scripts about sexuality to describe these men's experience creates a specific crisis of meaning-making. We will also illustrate how, in the absence of adequate social narratives, more highly developed cognitive structures can enable individuals to construct relatively adequate personal narratives.

» METHODS

Our analysis is based on a reading of messages posted by gay/bisexual men in self-described "mixed-orientation marriages" (MOMs) to three internet support groups between 2005 and 2010. These groups were a) HOW (Husbands Out to Wives) a group for men whose wives know of their same-sex desires; b) HUGS (Hope-Understanding-Growth-Support), a yahoo-based group for both members of mixed-orientation couples; and c) Closed Loop, a yahoo-based group for men in heterosexual marriages who seek exclusive sexual relationships with other men. Our method is participant-observation; we originally gained access to these posts not as researchers, but as members of the (private) listservs. Because it would be impossible to obtain informed consent from individuals to use posts as research data, we limit our analysis to a) a count of simple "factual" information to evaluate whether the men who wrote these posts present a profile similar to the participants in past studies of mixed-orientation marriages, and b) a qualitative description of our observations as participants in the group conversations on the lists over a five-year period. For the same reason, we do not quote passages from the posts.

We selected posts by men because the number of posts by gay/bisexual women in our sample was too small to make generalizations. Since participants in these groups might write frequently over several years, on multiple groups, using different names and email addresses, we used the individual message rather than the person as our unit of analysis. We eliminated posts that were merely indicating agreement with previous poses (e.g., "Right on!"), as well as those that did not address the topic of mixed-orientation marriage (e.g., making logistical arrangements for a meeting, recommending a book). This yielded 659 posts.

» RESULTS AND ANALYSIS

The crisis of the SSA spouse in a heterosexual marriage

Most research on men who find themselves in "mixed orientation marriages" in adulthood reports that they usually entered the relationship desiring and expecting to be satisfied with a traditional, monogamous heterosexual marriage. They typically had some homosexual experiences in adolescence or early adulthood, but believed that their same-sex desires were "a phase." They expected to be able to put them aside with marriage, just as two straight people entering marriage might expect to put aside their extramarital heterosexual attractions (Edser & Shea, 2002; Matteson, 1985; Ross, 1979). This pattern was true for all but a few of the men in our sample; only eight posts reported having sex with men during the first year of the marriage.

Past studies report that for most men, this construction "works" for many years. Frequently, it is only after many years of marriage that same-sex desires reach a level of intensity and significance that leads the heterosexually married gay/bisexual husband to re-examine this life-plan (Edser & Shea, 2002; Ross, 1979). Meanwhile, they have developed adult life patterns of intimacy and generativity in terms of the dominant cultural script for heterosexual marriage, including their marriage, children, relations to extended family and friends. Our sample supports this finding: in 83% of posts describing the period before the "crisis of coming out," there was a period of several years in which same-sex desire was not problematic. Nothing in our sample suggested that these men had any lesser commitment to and satisfaction with their "heterosexual lives" than other heterosexually married midlife adults.

At some point, however, same-sex desire re-emerged as a conscious concern and became motivating enough to cause significant conflict/unhappiness. Participants in the online groups identified several changes that seemed to facilitate this "return of the repressed." The most common attribution was to broader social changes regarding homosexuality (mentioned 72 times)—greater social acceptance, the visibility of gay people in media and social life, and, especially, the rise of internet sites which made it possible for individuals to access gay pornography, chat rooms, discussion groups, and social networks to facilitate arranging sexual encounters, all without leaving their homes. Other factors often mentioned included a lessening intensity of sex and romance in the marriage (55 posts), and a lessening preoccupation with the goals that had been central in early adulthood. These goals included children, career building, establishing a home (14 posts). Several men mentioned their own illness or the death or illness of parents as increasing their sense of mortality and intensifying the urgency of the need to deal with their unrealized desires (23 posts).

Recognizing their same-sex desires and disclosing them to their spouses (intentionally or accidentally, by getting "caught" in extramarital sexual activity), these men face the difficulties of coming out as homosexual in a homophobic culture and the marital crisis of infidelity (real or desired). They also face the identity crisis of having aspects of the self that feel vital to one's authenticity and wholeness. However, these men are in conflict with a structure of meaning and commitment that has defined one's earlier adulthood. Concretely, these men (and their wives) must decide whether to remain married. Buxton (1994) estimates that about 2/3 divorce relatively quickly, while a third decide to try to maintain the marriage. Of these, about half are still together five years later. Whether they ultimately remain married or divorce, however, these men face the challenge of reformulating an identity (and a life) that in some way addresses the meanings of both their heterosexual life and their homosexual desires.

As reported in another analysis of messages on MOM groups (Klein and Schwartz, 2001), most of the posts in our sample dealt with what might be called "practical" conflicts. These conflicts included how to renegotiate major relationships in the light of one's same-sex desires, coming out to friends and family, whether to maintain or end the marriage, the ethics of consensual and/or hidden extramarital sex dominated a good deal of discussion. In addition to these concrete difficulties, however, we observed that for many men the re-emergence of same-sex desire posed an unsolvable cognitive problem that created a crisis of meaning-making. The central predicament for many SSA spouses seemed to hinge on the ways in which their feelings, motives, and experiences were not adequately described by dominant cultural scripts about sexual orientation. In the attempt to renegotiate an identity that could make sense of both the "straight" and the "gay" parts of their lives, these men asked questions to which our dominant cultural understandings of desire, love, and marriage provide no answers or, alternately, multiple contradictory answers. In the absence of a sensible answer to the question "Who am I, sexually?" it was impossible to find purchase on the pressing questions regarding life decisions. Below, we examine the dominant cultural narratives regarding sexuality and how they fail men in mixed orientation marriages.

The "orientation script": an essentialist view of sexual desire

The narrative about sexual desire held most widely, both by psychologists and by the general public, is essentialism, the belief that people do not merely have differing sexual preferences and behaviors, but that they have different underlying *orientations*. For an argument that this is a historically recent view of sexual diversity, see Halperin, 1990, and Foucault, 1990; for an argument that sexual orientation is a socially constructed category, not reflective of a natural category, see Stein, 1999. The core assumptions of the essentialist view are:

1. People have a sexual orientation, an inner nature that leads them to desire either men or women (or in more sophisticated versions of the script, sometimes both). In other words, people come in one of two sexual kinds: gay or straight.

2. A person's sexual orientation is present at birth, probably rooted in biology, and does not change across his or her lifetime, even if his or her sexual behavior or his experienced desire does change.

3. All forms of sexual desire and romantic love reflect a person's sexual orientation (and thus will be directed at a single gender).

There are two major variants of the orientation script in modern American popular culture. The "heteronormative" version of the orientation script assigns positive value and health to a heterosex-

ual orientation and sees homosexual orientation as sinful and/or pathological. The "gay affirmative/identity" version of the script argues that the two orientations are equally normal/healthy. This identity-affirmative version sometimes adds one more assumption to the orientation script, namely,

4. People should adopt a private and public identity that reflects their sexual orientation. They should act on it by having sexual and romantic relations with the gender they are attracted to.

The orientation script assumes, then, that sexual preferences are not like, say, tastes in food or political beliefs. A person can in general find spaghetti his or her most preferred dish, but enjoy the variety of an occasional lobster dinner, or be liberal in one's youth and become more conservative with age, without contradiction—not so with regards to love, sex, and marriage. The script takes one's sexual orientation to be a fundamental fact of one's being, a core aspect of identity. It defines who a person is and how he is socially situated. It also assumes that "the whole package"—lust toward strangers, sexual pleasure in relationships, romantic love, commitment and family—will all be directed at a particular gender, providing a basis for a coherent sexual/affectional life that is either gay or straight.

How the "orientation script" problematizes the SSA spouse's identity

The essentialist script was the dominant cultural construction of sexual identity when the men in our sample were adolescents and young adults (roughly, 1970–2000). Moreover, the dominant version of the script was binary (gay vs. straight are the only categories) and heteronormative (heterosexuality is more valued). Almost all of the men reported experiencing desire for other men in their adolescence, most had homosexual encounters, and a few had romantic relationships with men prior to marriage. When they fell in love with a woman, and in that context enjoyed heterosexual sex, the essentialist script offered them a choice. Their sexual identities and adult lives could be based around one or the other of their desires. Most were relieved to be able to escape the stigmatized category "gay" and to be able to lead a "normal" heterosexual life. Their development of sexual identity is represented below.

When the press to recognize and act on homoerotic desire became more intense, in later adulthood, most of these men continued to try to make sense of their feelings in terms of the orientation script—despite the fact that their experience contradicted it. Their sexual desire and romantic love were **not** consistently directed toward one gender, nor did they usually feel similar kinds of love and desire toward men and women. It is evidence of the power of social scripts that SSA spouses rarely concluded that the orientation script must be wrong. Instead, they struggled, often desperately, to find a way to make sense of their feelings in terms of the dominant cultural narrative, asking them, "*What am I? Am I straight? Then why do I desire sex with men so intensely? Am I gay? Then why do I love and desire my wife? Am I bisexual (if that even exists)? Then why are my sexual fantasies only about men?*" It is not surprising, then, that "labels" (for sexual orientation) are a topic of regular and heated discussion on internet groups for men in MOMs (Klein & Schwartz, 2001).

The SSA spouse's understanding of his feelings for his wife and for men, his evaluation of his past choices and the current meaning of his marriage and his desire for a homoerotic life all are shaped by whether he sees himself as "really gay" or "really straight," or perhaps "really bi." The choice of a label delimits what options are available to him and defines their meaning and value. Many men attempt to make sense of their homoerotic experiences in terms of the gay-identity script and draw upon the positive narratives of the gay community to give shape to, legitimize and support recognition and acceptance of same-sex desire. From this point of view, the SSA spouse has a right and perhaps a moral imperative to act on the same-sex desires, to be "true to himself" and live "authentically." Taking this perspective, the SSA spouse may feel that his wife has an obligation to support his expression of gay desire, that to do otherwise would be selfish and uncaring on her part. However, these narratives delegitimize and problematize his marriage and his love and desire for his spouse, seeing them as "mistakes," made, perhaps, out of "internalized homophobia." Conversely, the heteronormative script delegitimizes and problematizes his same-sex desire. What both scripts agree on is that the two do not or should not coexist—not in one person or in one life.

experiences/motives	interpreted via essentialist script	identity/life choices
I feel sexual desire for men/enjoy sex with men, AND I fell in love with a woman/enjoy sex with a woman.	experiences with men not integrated into identity; life choices based on "real" (heterosexual) orientation.	Only one of these can be "true/real." The other must be "a phase," not really me.

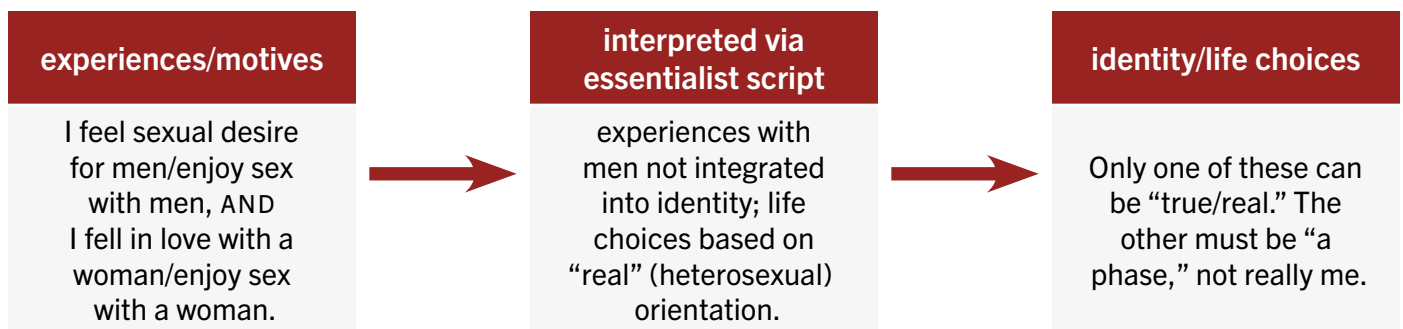**Figure 1.** t The development of sexual identity in men in our sample

Table 1. implications of sexual scripts for SSA spouse's identity and life

	identity	life choices
hetero-normative	Continue efforts to exclude/repress same-sex desires in the face of contradictory experience or accept same-sex desire and reconstruct previous identity as "false." Replay of adolescent identity crisis in the sexual sphere.	Heterosexually-based life choices (marriage, family) often continued, but experienced as "false" or "a sham." Can lead to despair, sense of a wasted life.
gay identity	Same-sex desires accepted and integrated into (revised, gay) identity; this requires excluding/repressing opposite-sex desires, often reconstructing them as "false."	Heterosexually-based life choices (marriage, family) now seen as "a mistake." Either discontinued and a new "gay" life begun, or revised to not conflict with new gay identity.
queer	Seeming contradiction between same-sex and opposite-sex desires transcended by relating them in a meta-construction of sexual identity. Both seen as elements of a larger system of sexual tastes and preferences. Stability of identity is located at the level of the system, particular desires are not seen as indicators of the whole.	Heterosexually-based life choices (marriage, family) seen as part of a larger process of acting on sexual identity across one's life. Modification of life choices is based on both same-sex and opposite-sex desires, as well as other sexual, romantic, moral, and practical considerations.

Some SSA men in MOMs find that the sexual orientation script provides an adequate description of their experience and a basis for their identity; they are able to redefine themselves as "gay" or to reaffirm their identity as "straight." Both of these paths, alternate ways of fitting into a socially normal script, offer a coherent meaning system and an affirming community. If, for example, a man has experienced relatively little romantic love and desire for his wife, in contrast to the intensity of his love and desire for men, he may decide that he was really gay all along. He may decide that he deceived himself to a degree about his love and desire for his wife because of his fear of living a public gay life. Viewed in this light, there will be a moral imperative for him to live authentically, as his "real self." The gay community will offer support for this choice, provide opportunities for new relationships and a form of life that will give meaning to his same-sex desires. "Coming out," to himself and to others, as "gay" may enable a joyful re-appropriation of the repressed parts of the self and a more inclusive and coherent basis for adult identity and life.

Alternately, a man may find the strength of his commitments to his heterosexual identity (deriving from, e.g., love for his wife and children, traditional social and religious values) lead him to continue to exclude his homoerotic desires from his core identity. He can affirm the meaning in his choice to marry, his history with his wife and its place in a network of straight relationships (children, extended family, straight friends, perhaps religious communities). From this perspective, the same-sex desires become a quirk, something that he needs to manage, like an addiction, to keep it from destroying his core identity, values, and commitments.

There is some value in resolving the conflict among the competing and incoherent parts of being an SSA spouse by adopting one or the other of these solutions and accepting an identity and social role as either gay or as straight and monogamously married. These solutions allow people to make the pressing choices—act on same-sex desires or not, stay married or divorce—with conviction and get on with life. Most of the men who participate in the groups we studied, however, find both of these alternatives unsatisfactory. For these men, both solutions—to stay married and minimize the significance and disruption of the gay desires, or take on a

gay identity and dis-identify with the marriage—exact a price in wholeness and genuineness. Both require the SSA spouse to disown and devalue some part of his experience, to conform to culturally defined forms of desire and to feel shame about the desires that are not culturally sanctioned. In short, while our culture offers scripts for gay sex/love/ marriage/identity, and scripts for straight sex/love/marriage/identity, it offers no coherent and meaningful way to integrate love and commitment to an other-sex spouse with significant same-sex desire. There are not, for example, generally available and adequate descriptions of the sexual orientation of someone who simultaneously experiences love and desire for two differently sexed persons, nor models of lives in which both of these desires are lived out in meaningful, positive ways.

"Queer" conceptions of sexuality and the role of adult cognitive development

If the SSA married man is to develop an identity and a life that integrates his various sexual and romantic motives, he must develop a conception of sexual desire and sexual identity. This desire and identity unseats the assumptions of the orientation script and resolves the contradictions that result from applying it to his experience. We did observe, in a small number of posts, such alternate conceptions of the SSA spouse's sexuality and explorations of the "non-normal" life choices they made conceivable. (These posts were more frequent on the internet groups for people in marriages in which the same-sex desires were known to both partners). We generally characterize these sexual scripts as "queer," in the sense that they do not accept the assumptions of the dominant models of "normal" sexuality (Warner, 2000). Queer conceptions of sexuality do not make presumptions about the aspects of other people that might be sexualized for an individual (e.g., gender, personality, hair color, race, SES), nor about the ways that different sexualized elements might be organized (in narratives about power, nurturance, exposure or hiding, pain, etc.). In addition, from a queer perspective, there is no a priori reason to expect people's organization of sexual motives to remain stable across their lives. In short, from an essentialist perspective the identity question is "What are you?" (meaning "which gender do

Table 2. ways that dialectical thinking supports the creation of a queer conception of sexuality

dialectical schemata	implications for construction of sexual orientation
thesis-antithesis-synthesis movement in thought.	orientation toward finding construction which includes both same- and other-sex desires in identity.
affirmation of the primacy of motion.	changes in sexual desire across lifespan can be included in identity.
avoidance or exposure of objectification, hypostatization, and reification.	recognition that "sexual orientation" is an abstraction, that motives and behavior are the primary reality.
understanding events or situations as moments (of development) of a process.	ability to integrate differences in sexual motives at different points in lifespan into a continuous narrative.
location of an element (or phenomenon) within the whole(s) of which it is a part.	specific desires and behaviors seen as part of larger system that organizes multiple motives.
understanding the resolution of disequilibrium or contradiction in terms of a notion of transformation in developmental direction.	conflicts among sexual motives/disequilibrium of identity (and life structure) can be seen as providing a potential for growth; integration can be valued without devaluing any elements.

you prefer?"). From a queer perspective it is "What is the (current) shape of your desire? What sexual acts and objects attract you, in what ways, and when?" Table 1 summarizes the three models of sexual orientation and their implications for the identities and life-choices of SSA married men.

To create a queer conception of sexual identity, the SSA spouse must reflect on his essentialist beliefs, understand them to be constructions rather than realities, compare them to his experience, and respond to the contradictions that result by formulating more complex constructions that can bring into relation the contradictory elements. We suggest that these kinds of moves in thought reflect capacities described in models of adult cognitive development (e.g., Commons, Richards & Armon, 1984; Basseches, 1984; Kegan, 1994). The methodology of this study does not enable a systematic analysis of the ways in which SSA men used advanced cognitive structures in resolving their identity conflicts. We can comment, however, on several patterns of thought we saw in the posts that illustrate the ways in which greater cognitive complexity supports the creation of a queer conception of sexuality and thus makes possible a more adequate integration of parts of the sexual and relational self.

Dialectical thinking, as described by Basseches (1984), orients thought toward perceiving processes of change, and recognizing and integrating contradictory elements into higher-order structures. Table 2 gives some examples of ways that dialectical moves in thought could support the creation of a queer understanding of sexuality. The dialectical attention to conflicting elements promotes the simultaneous consideration of both aspects of the gay/bi spouse's sexuality and discourages restoring harmony by "suppressing" one or the other. The dialectical tendency to see

change rather than fixity supports viewing one's sexual identity as in flux rather than as static across the lifespan. Higher order dialectical schema (those which integrate understanding of change/process and understanding of structure) provide a framework through which the gay/bi spouse can reflect upon his individual motives and experiences as parts of larger wholes, and bring those wholes into relation in hierarchically more differentiated and integrated structures—in other words, to understand one's "straight" and "gay" experiences as embedded in different cultural scripts, and to integrate those conflicting scripts in a higher-order "queer" perspective.

Similarly, Kegan's "fourth-order thinking" describes cognitive moves that could enable gay/bi spouses to include both their same sex attraction and their love for the wives in their identities and life choices (see Table 3). In third-order thinking, according to Kegan, meanings are validated by external sources—other people or groups. Thus, the individual understands both the meaning of his marriage and the nature and meaning of his same-sex desire in terms of the very cultural scripts and expectations that we have argued create a conflict of "incommensurability." Fourth-order thought, in contrast, opens the possibility of self-created meaning systems which include and relate particular relationships and social scripts, but ultimately locate their validity in the authority of the self. Fourth-order thought allows gay/bi spouses to create definitions of marriage and sexual identity that are not socially affirmed, and to recast traditional scripts so as to reduce dissonance. (Kegan also describes "Fifth-order thinking" consisting of meta-systematic moves in thought; this form of reasoning is probably necessary for the full development of a "queer" conception of sexuality and relationships, but was rarely seen in our sample).

Table 3. ways that men conceptualized issues related to their MOM's using 3rd and 4th order thinking

issue	3rd order thinking	4th order thinking
future of marriage	depends on whether it can be put "right" again, whether traditional norms and roles can be successfully reinstated.	view of relationship as process, success is not predictable, depends on evolution. often asserts that the instability of an mom is not fundamentally different from any marriage
understanding of sexual orientation	SSA partners' homoerotic desire is "who they are" and thus contrary to participating in their marriage.	SSA partners' homoerotic desire is "something they have," a relational role among many, and they can choose how to relate it to their role as marriage partner.
the past marriage	was false, not real, a mistake. straight spouse's perception of the relationship was mistaken.	what was perceived was real, but now is understood differently in the larger context created by understanding something which was hidden/distorted.

» SUMMARY

The combination of an extended lifespan and the rapid social change characteristic of post-modern culture is likely to put more mid-life adults in situations where parts of the self that were "left behind" in adolescence can re-emerge with increased intensity. The men discussed in this paper illustrate not only of the difficulties of renegotiating identity in adulthood, but the particular predicament of people whose significant meanings and experiences cannot be held in meaningful relation by current cultural scripts. Those who attempted to resolve their identity and life problems within cultural scripts for sexual orientation were forced to sacrifice, as they had in adolescence, some aspect of their sexual experience to achieve coherence and connect to a supportive community. Others were able to recognize and accept the reality of their experience, including the parts that conflicted with cultural scripts, and to create identities that integrated previously disparate parts of the self and imagine a range of life choices—both "normal" and "un-normal"—that were not based in sacrificing important meanings. Our analysis of the writings of men who took each path suggests that advanced adult cognitive operations may be required to negotiate such "culture-transcending" renegotiations of identity. ■

REFERENCES

Basseches, M. (1984). *Dialectical thinking and adult development.* New York: Ablex Publishing.

Buxton, Amity Pierce (1994). *The other side of the closet: The coming-out crisis for straight spouses and families*. New York: John Wiley & Sons.

Commons, M. L., Richards, F. A. & Armon, C. (Eds.). (1984). *Beyond formal operations: Vol. I: Late adolescent and adult cognitive development*. New York: Praeger.

Erikson, E. H. (1968*). Identity: Youth and crisis*. Oxford:Norton & Co.

Foucault, M. (1990). *The history of sexuality: An introduction*. New York: Random House.

Edser, S. J. & Shea, J. D. (2002). An exploratory investigation of bisexual men in monogamous, heterosexual marriages. *Journal of Bisexuality*, *2* (4), 7–39.

Halperin, D. M. (1990). *One hundred years of homosexuality*. New York: Routledge.

Kegan, R. (1994). *In over our head*s: *The mental demands of modern life*. Cambridge, MA: Harvard University Press

Klein, F. & Schwartz, T. (Eds.). (2001). *Bisexual and gay husbands: Their stories, their words*. New York: Haworth Press.

Marcia, J. E. (1994.) The empirical study of ego identity. In H. A. Bosma, T. Graafsma, & L. G. Tobi (Eds), *Sage focus editions: Vol. 172. Identity and development: An interdisciplinary approach*. (pp.67–80). Thousand Oaks, CA: Sage.

Matteson, D. R. (1985). Bisexual men in marriage: Is a positive homosexual identity and stable marriage possible? *Journal of Homosexuality, 11*(1–2), 149–171.

Ross, M. W. (1979). Heterosexual marriage of homosexual males: Some associated factors. *Journal of Sex and Marital Therapy, 5* (2), 142–151.

Stein, E. (1999). *The mismeasure of desire: The science, theory, and ethics of sexual orientation*. New York: Oxford University Press.

Swan, T. B. & Benack, S. (2012) Renegotiating identity in unscripted territory: The predicament of queer men in heterosexual marriages, *Journal of GLBT Family Studies, 8* (1), 46–66.

Warner, M. (2000). *The trouble with normal: Sex, politics, and the ethics of queer Life*. Cambridge: Harvard University Press.

Westen, D. (1985). *Self and society: Narcissism, collectivism and the development of morals*. New York, Cambridge University Press.

BEHAVIORAL DEVELOPMENT BULLETIN

An analysis of life interviews selected for ratings of life satisfaction correlated with ratings of dominance

Nancy Nordmann

National Louis University

ABSTRACT

Historically collected data and the analysis of that data are presented demonstrating the use of seminal measures of life satisfaction and interpersonal behavior in identifying a model of interpersonal functioning, autogenesis, the elaboration of which addresses an issue of contemporary focus, the integration of the interpersonal constructs of agency and communion. These seminal measures are demonstrated to be of continuing value as Index Measures of autogenesis and to provide criterion validation of the autogenetic model.

Autogenesis is proposed as an interpersonal dynamic defined as the development of the potential of a sense of origin, energy, self-direction, effectiveness and ultimately communion in one's activities as they relate to the world of others. Autogenesis is posited as expressing, through a number of thematic categories of interpersonal perception and functioning, the growth of the self as an ecologically adapted origin and agent. The suggested goal expressed by the autogenetic process is that of a self-directing, self-maintaining, self-realizing entity that is responsive in an interpersonally structured world.

Hypotheses examined are: the reliability of identifying autogenetic stages in individual life interviews; the validity of identifying autogenetic stages using an Index Measure; and the criterion validity of the autogenetic model.

KEYWORDS: interpersonal agency, interpersonal development, agency and communion

T HIS PAPER REPORTS PREVIOUSLY unpublished results of an analysis completed in 1991 of data collected between 1974 and 1979 identifying perspectives of interpersonal agency denoted as autogenesis. The suggestion is that historical data and analyses can be revisited as a source of relevant contemporary insights.

In 1975 Morton Lieberman and Bertram Cohler of the University of Chicago completed The Ethnic Study of personality correlates of life satisfaction. A sample of 285 first and second generation Chicago Irish, Italian, and Polish men and women, 40 to 80 years of age were given the 20 item Life Satisfaction Index (LSI) developed by Bernice Neugarten (1961) of the University of Chicago and 19 personality measures in use at the time, one of which was a 48 item Interpersonal Adjective Checklist (ICL) developed by Timothy Leary (1957) scored for Dominance (ICL-D). The ICL-D of the personality indices used in the study proved to be the best predictor of life satisfaction, $r = .29$, $p < .001$.

Author note: Nancy Nordmann, School of Social and Behavioral Sciences, National Louis University, Chicago, Illinois. This article is based on Nancy Nordmann's doctoral dissertation at the University of Chicago. Work on this article was supported in part by a sabbatical appointment as a Visiting Scholar at the Harvard Graduate School of Education in Human Development and Psychology. Correspondence concerning this article should be addressed to Nancy Nordmann at nnordmann@nl.edu or nnordmann@uchicago.edu.

Correlations of .30 represent a moderate positive effect size according to the Cohen (1988) conventions to interpret effect size. Correlations in the measurement of personality generally fall in the range of .30, which although typical in psychological research is often viewed as problematic when compared to the Cohen range of above .50 for a large effect size. Hemphill (2003) derived "empirical guidelines concerning the magnitude of correlation coefficients found among psychological studies" (p .78) extending Cohen's benchmarks. Hemphill interpreted magnitudes of correlation coefficients of .20 to .30 as representing the middle third of the distribution and greater than .30 as representing the upper third. Hattie (Lenhard & Lenhard, 2014) describes .29 correlations in educational research as falling near the middle of the zone of desired effects. The Ethnic Study correlation found between ratings of life satisfaction and dominance led to an investigation of this relationship, which seemed merited at the time and to an even greater extent currently, considering contemporary effect size guidelines.

In order to investigate the correlation between ratings of life satisfaction and dominance found in the Ethnic Study, 27 of the male subjects were chosen for an open-ended life interview. The interviews obtained were recorded and transcribed and copies

given to the author to analyze. The resulting analysis (Nordmann, 1991) suggests that dominance as measured by the ICL-D reflects various elaborations of interpersonal agency, denoted as autogenesis, and that positive and negative Life Satisfaction Index ratings within each of the elaborations of autogenesis reflect a dialectic that contributes to shifts within and between elaborations.

Ethnic study measures that reveal autogenetic functioning

Neugarten life satisfaction index (LSI). Neugarten represents those survey researchers after World War II "who began polling people about their happiness and life satisfaction using simple global survey questionnaires" (Diener, Oishi & Lucas, 2009, p. 189), of which there are now many. Subjective measures of satisfaction are now being adopted not only by social science researchers, but also by governments (Diener, Englehardt, & Tay, 2013) and Non-Governmental Agencies (Pew Research Center, 2014; OECD Better Life Index).

In their review of research on subjective well-being, Diener, Oishi and Lucas (2009) present a contemporary example of a life satisfaction measure, the five-item Satisfaction with Life Scale. Three of the five items are near identical to LSI items. "I am satisfied with life. So far I have gotten the important things I want in life. If I could live my life over, I would change almost nothing" (Diener, Emmons, Larsen & Griffin, 1985). The LSI further elaborates these items. Both the LSI and the Satisfaction with Life Scale have been used extensively and exhibit high degrees of stability (Adams, 1969; Diener, Inglehart & Tay, 2013, p. 499). The similarity in content, extent of use and degree of stability of these measures suggest the continuing relevance of the LSI as a measure of life satisfaction.

As an aside, Neugarten chaired the Committee on Human Development at the University of Chicago, as did Morton Lieberman. Although departments and programs of Human Development are now ubiquitous, the institutionalization of human development as an academic discipline had its beginning in this committee at the University of Chicago. That Neugarten's early instrument and Lieberman's and Cohler's investigation might continue to bear fruit lends additional historical significance to an auspicious beginning of the organized study of human development; a beginning oriented around a perspective that presaged today's field of positive psychology. Mihaly Csiksentmihalyi a progenitor of the field of positive psychology began his professorial career in the Committee on Human Development and there conducted research that resulted in the study of the optimal state of intrinsic motivation reported in the book *Flow* (Csiksentmihalyi, 1990). The psychology of happiness is one of the more recognized recent disciplines within positive psychology. And as indicated above, subjective survey measures of happiness continue in the tradition of the LSI.

Leary interpersonal adjective checklist scored for dominance (ICL-D). Timothy Leary who is perhaps best known for his promotion of experimentation with psychedelic drugs, encouraging youth of the 60's to 'tune in, turn on and drop out' and whose ashes were rocketed into an orbit around the earth in 1997, produced credible work in personality measurement at Kaiser Permanente Hospital in Oakland as director of clinical research and psychology. The work of the Kaiser team resulted in the monograph, *The Interpersonal Diagnosis of Personality* authored by Leary (1957)

that was awarded the best book on psychotherapy in that year (Strack, 1996). At the 1994 American Psychological Association convention Leary was recognized at a symposium honoring his work in interpersonal psychology titled "Interpersonal Theory and the Interpersonal Circumplex: Timothy Leary's Legacy". The Journal of Personality Assessment published a special series by the same name in 1996.

Leary (1996) approached the clinical assessment of personality in a humanistic and person-centered way within the context of interpersonal transactions as proposed by the psychiatrist Harry Stack Sullivan (1953). Leary cites the psychologists Maslow, McClelland, Murray and Rogers as influences emphasizing the positive growth potential inherent in human beings. Focusing on interpersonal interaction and potential, Leary and his colleagues developed an interpersonal circumplex model of personality. Of Leary's colleagues, Rolfe LaForge of the Kaiser Foundation Hospital and Robert Sugzek of the Permanente Psychiatric Group (LaForge and Suczek, 1955) best provide a description of the circumplex model.

The data of interpersonal behavior from each of the levels of personality are ordered in terms of a classificatory system made up of 16 basic interpersonal variables. These are arranged in the form of a circular continuum defining the relationships between elements, i.e., the theoretical degree of relationship between any two variables is a decreasing function of their separation on the perimeter of the circle. Thus, variables juxtaposed on the perimeter of the circle are theoretically similar and should be highly correlated while variables on opposite sides of the circle are logically opposite and should be negatively correlated. A varying degree or intensity of any one of the 16 variables can also be represented in the circular schema by the distance at which it is placed (or scored) along the radius from the center of the circle. Thus traits represented nearer the center are considered to be of normal, moderate, or appropriate intensity while those at the circumference are considered to indicate an abnormal degree or intensity of the same trait. (p. 96)

This circular model permits dimensionalizing of an individual's traits. This approach differs from the approach of the Diagnostic and Statistical Manual of the American Psychiatric Association (DSM) that categorizes and labels individuals discontinuously according to symptoms. In the circumplex model a person has more or less of each interpersonal characteristic rather than an assigned category or label. This dimensionalizing rather than categorizing distinguishes a humanistic from a medicalized approach to assessing personality.

An International Journal of Testing review of the assessment of personality using adjective check lists (Craig, 1995, p. 178) reports a long history of adjective checklist methodology beginning with Hartshorne and May, who in 1930 developed a list of 160 words consisting of 80 pairs of antonyms related to four types of conduct, and Alport and Odbert, who in 1936 surveyed English dictionaries for adjectives used to describe personality. The Leary Interpersonal Adjective Check List is listed on the Current Adjective Check List Personality Tests in Print Table 1 (Craig, 1995, p. 180) and continues to be available. The model, theory and validity and reliability of
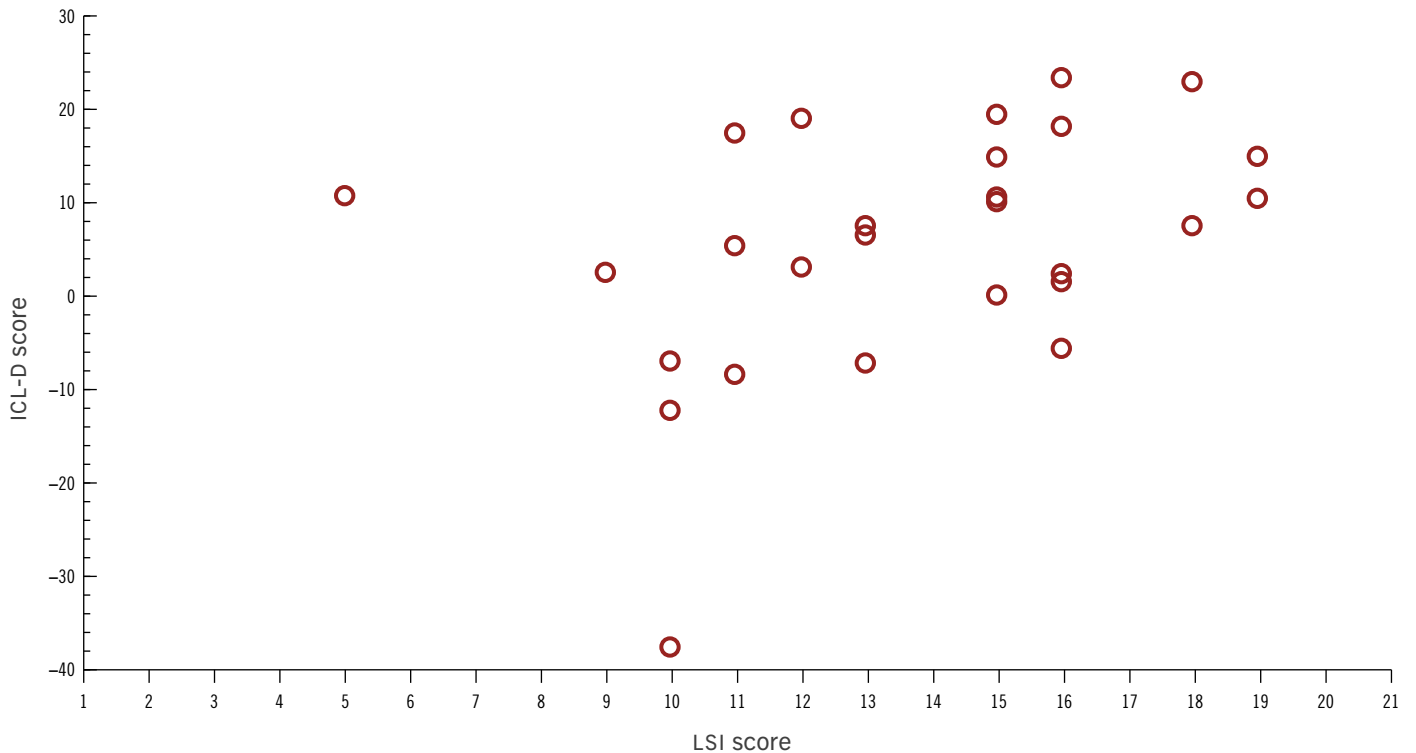
Figure 1. plot of interview subjects' LSI and ICL-D scores

the measure are detailed in the Leary monograph (1957). Acton (2002) notes "The validity of the ICL scales haw been investigated in more than 300 studies (summarized by Taulabee & Clark, 1982). Among these, Leary's early studies (e.g., Leary, 1957) remain noteworthy validations" (p. 449).

Locke (2011) acknowledges the ICL as the first Interpersonal Construct (IPC) measure; it's use in numerous studies; and since, the construction of interpersonal scales for a variety of domains. One of these is the Interpersonal Adjective Scales (IAS) (Wiggins, 1995; Wiggins, Trapnell, & Phillips, 1988). Lowe reports the psychometric and circumplex properties of the IAS to be superior to those of the ICL and that the IAS is now the preferred measure of interpersonal dispositions. This outcome does not negate an analysis utilizing the ICL, the resulting model of interpersonal perspectives or the renewal of the value and use of the ICL for identifying these interpersonal perspectives.

Analysis of the ethnic study life interviews

The research undertaken by the author involved analyzing life interviews of 27 subjects of the Lieberman and Cohler (1975) ethnic study of life satisfaction and personality; seeking the nature of the factor contributing to the association of life satisfaction with the ICL scored for Dominance.

Method

Free form interview data was collected for 27 of the 285 Ethnic Study subjects, 14 with high and 13 with low life satisfaction ratings. David Gutmann of the Committee of Human Development at the University of Chicago interviewed each subject for three hours regarding themselves and their lives. The transcribed interviews were

made available to the author to analyze for life patterns that might be associated with high and low life satisfaction. LSI and ICL-D scores for the Ethnic Study Interview sample and LSI and ICL-D scores for a Sage Study Self-help sample were also made available.

Samples. The Interview (I) sample was comprised of 27 males chosen to fill cells generated by the characteristics of age (40–55, 55–70), ethnic affiliation (Irish, Italian, Polish) and Life Satisfaction (high or low). The mean score for LSI in the original ethnic sample was 13.42. The possible scores ranged from 1–20. The mean on LSI for the interview (I) sample was 13.64. The median in both cases was 14. The assumption was that the I sample, though composed of males only, and was normally distributed on the LSI. The I sample was also normally distributed on the ICL. The I sample ICL scores ranged from −38 to + 24. The original sample ICL mean was 4.51. The I sample mean was 5.58.

Prior to the initiation of the interview analysis, a second sample for which there were LSI and ICL-D scores became available. The second sample was comprised of 15 males drawn from a group of 87 males and females aged 60–80. These individuals lived in San Francisco and had sought help for their interpersonal problems at a self-help center called Sage. As a part of their Sage activities these individuals volunteered to take part in an evaluation program linked to their Sage experience. The LSI and ICL were among the instruments used in the evaluation.

Procedure. An analysis of the ICL items contributing to the dominance score suggested that the factor being tapped was not 'dominance' in the generally accepted sense of ruling or ascendancy. Descriptions of being in charge, a good leader, and dominating contribute to the score, but so do descriptions of tenderness, leniency and taking care of and spoiling others.

Table 1. interpersonal engagement of interview subjects rated from low to high on a combination of LSI and ICL-D

dependent	
submerged sense of individuality	nearly unrelated to others
	acceptance and complacency with regard to others' control
	a sense of necessity and obligation in interpersonal relations on the given individual's part
	rebellion and escapism
transitional	
emergent sense of individuality	self-limitation with regard to others
	the selection of a protected environment, limiting the environment
	expectation of and manipulation of others
	opportunism in relationships
	effortful participation in partnerships
independent	
emerged sense of individuality	active management of balance in partnerships in addition to effort in partnerships
	a sense of independent management failed
	independence, concern with dependents, limiting their activities
	independence, concern with dependents, limiting oneself with regard to what one will do vis-à-vis dependents
	personal balance and momentum and interpersonal responsiveness

The first step of the analysis was to plot the subject' life satisfaction and interpersonal scores. The observable pattern of regression supported the assumption of the correlation of these two variables; see Figure 1.

The second step was to group the interviews according to life satisfaction and examine the groups for similarities and differences. The interviews were divided into two groups, above

and below the life satisfaction mean, $M = 13.66$, and read by group. The interviews were then grouped according to both the life satisfaction and interpersonal mean, $M = 5.73$. This approach resulted in four groups of interviews based on ratings of above or below the mean of the two measures: a) low satisfaction/low interpersonal; b) low satisfaction/high interpersonal; c) high satisfaction/high interpersonal; and d) high satisfaction/low interpersonal. The interviews were read seeking within group similarities and between group differences.

General distinctions in interpersonal functioning related to satisfaction and dissatisfaction began to emerge. Subjects with similar scores on both measures (14 clusters of scores) expressed themselves in similar and distinct ways. When these clusters of individuals were arranged from lowest to highest ratings and compared, three general groups of interpersonal engagement were identified; dependent, transitional and independent; see Table 1.

A number of strands seemed to unite the groups. From lower to higher level groups subjects seemed to shift in dependent to independent postures with regard to interpersonal involvement and agency. The subjects' perceived and actual involvements with others appeared to become more actively voluntary and effective. There appeared to be an increased sense of personal momentum and self-initiated, personally controlled, directed, and maintained effort that was of an adaptive value both to the individual and those around him.

The strands that seemed to unite the groups exhibited a dynamic quality that occurred across all the groups. This unifying dynamic was labeled autogenesis. Autogenesis, or personally controlled momentum, as judged by effectiveness, seemed best to account for the differences distinguishing each of the groups. The various expressions of interpersonal relations the groups represented appeared to be differentiations of this unifying dynamic.

Autogenesis, this sense of personally derived and personally maintained momentum and effectiveness seemed to increase sequentially, in a stepwise fashion, across the groups. The sequence of differentiations was similar to stages identified by developmental stage models of social development, suggesting that the 14 groups could represent developmentally related stages (Fowler, 1981; Kohlberg, 1964; Loevinger, 1966, 1976, 1984; Selman, 1980).

Autogenesis was seen as epigenetic development from an immersed, reactive self toward the emergence of an interpersonal self, the goal of which is self-balance, momentum, and the maintenance of effective, energetic, personal activity or involvement coupled with responsiveness to others. Autogenesis was seen as describing the natural growth of self-initiative, self-maintenance (psychologically), and self-realization in an interpersonal environment.

Development of an autogenesis coding manual

A coding manual was created that identified features and behaviors associated with interpersonal functioning for each of the originally observed fourteen autogenetic groups. It was hypothesized that the groups of interpersonal functioning described by the autogenetic manual, if valid, should be codifiable into a manual whereby a rater, other than the original researcher, could reliably identify a subjects' autogenetic group of interpersonal functioning.

The coding manual was used by an independent rater to code six subject interviews originally coded by the researcher in identifying the autogenetic groups. The rater did not code one of the interviews, deeming that nothing in the interview fit into any of the categories. She coded one interview at two stages contiguous to one another, one of the ratings being in agreement with the original rater and the other not.

In addition to rating each individual, the independent rater made notations regarding group descriptions. She sought to clarify any confusion she experienced with the manual. In instances where there was a discrepancy between raters, an analysis of the notations indicated the rating discrepancy was due to incorrect interpretation of the group descriptions. These misinterpretations were clarified and the clarifications incorporated into a revised version of the coding manual.

The researcher confirmed the clarification concerning the coding manual with the independent rater. The independent rater utilized the revised manual to rate an additional five interviews previously rated by the researcher.

Interrater reliability

The percentage agreement between the original rater and the independent rater was calculated, including and excluding the scorable interview and using each of the scores for the interview coded at two stages. The lowest percentage agreement obtained between the original rater and the independent rater yielded a

Table 2. autogenetic stages

1	uninvolved
2	accepting
3	rebellious
4	directed by a sense of necessity, obligation and responsibility
5	inclined to escape
6	prone to limit oneself or stint
7	inclined to choose a limited or defined and safe environment
8	manipulative
9	opportunistic
10	inclined to put forth a unilateral single handed effort in partnerships
11	inclined to manage balance in partnerships, to speak out
12	working within a complex system of individuals and needs and determining how the system functions
13	aware of the limitations or failures of the system of relations in which one functions
14	independent and concerned with establishing limits for dependents
15	independent and concerned with specifying one's activities regarding dependents
16	independent and responsive to others

Pearson *r* of .748, *p* < .01. The interrater reliability obtained between the independent rater and the original researcher utilizing the unscored interview, the double scored interview discrepant score, and the interviews for which the independent rater utilized the confirmed revised version of the coding manual yielded a Pearson *r* of .931, *p* < .001.

The high percentage of agreement between raters utilizing the coding manual and its revision suggests that the autogenetic coding manual or category rating system distinguishes interpersonal differences in interpersonal functioning between individuals in a reliable way from coder to coder.

Theoretical consistency

Following the test of interrater reliability, the autogenetic structure, the categories, and the coding manual were reviewed for theoretical consistency. Excluding the first and last groups, all of the remaining originally observed groups, with the exception of two, occurred in more or less antithetical pairs: necessity, obligation, responsibility / escapism; self-limitation / limiting the environs; manipulation / opportunism; silent, one-sided effort in partnerships / outspoken balanced effort in partnerships; and independent—limiting dependent's activities / independent—limiting dependent's activities regarding dependents.

The two original unpaired categories, other than the first and last, were the original group 2, 'accepting', and the original group 12 'system limitation or failure'. For theoretical consistency a paired group was theorized for each of the two originally unpaired groups. A 'rebellious' group was postulated and paired with the group of 'accepting'. A 'system functioning' group was postulated and paired with the group of 'system limitation or failure'. These groups embodied experientially observed characteristics unaccounted for in the original autogenetic sequence, were theoretically consistent with the structure of autogenesis, and filled gaps in the original sequence of clustering of the index LSI and ICL-D scores, which may have occurred simply as a result of incomplete sampling. The two theorized groups were interpolated into the autogenetic sequence of groups yielding 16 as opposed to 14 groups.

The first autogenetic group, 'uninvolved', is associated with the experience of a low level of life satisfaction. The last autogenetic stage, 'independent and responsive to others', or 'integrated' as opposed to 'uninvolved', is associated with a high level of life satisfaction. No experiential or theoretical considerations were apparent to recommend that 'uninvolvement' have a positive

life satisfaction expression, or expression, or that the successful attainment of 'independence and responsiveness' have a negative life satisfaction expression. In fact there exists, from a modeling perspective, a structural and logical symmetry in a solitary negative initial position in the organization of self-other relations and a solitary positive peak attainment position, between them, enclosing seven sets of progressively more differentiated pairs of self-other relations dialectically related, each expressing a negative followed by a positive experience of life satisfaction. For these theoretical reasons and for the lack of meaningful experiential evidence to the contrary, no antithetical categories were postulated to be paired with the first and last categories of the autogenetic sequence. The '16 classifications' autogenetic model was used in additional confirmatory research undertaken not reported here; see Table 2.

Index measure of autogenesis

The Index Measure of Autogenesis utilizes scores from the LSI and ICL-D. Plotting the subject's LSI score against the ICL-D score results in an autogenetic category identification. Scores on the LSI are placed in one of two categories, below the mean score for the original I sample, *M* = 13.66, or above the mean score. The range of ICL-D scores exhibited by the original I sample was divided at seven points as determined by the grouped scores yielding nine levels of ICL functioning. Thee nine ICL-D levels, combined with the two LSI categories, represented observed groupings of distinct and similar subject responses. The nine level parameters, expressed in ICL-D scores, represent the means of the mean scores of adjoining levels or groups of ICL-D scores. A subject's ICL-D score is plotted as being below the first parameter (−24.76,); above that parameter, between that parameter and the next parameter (9.51); or between −9.51 and −2.38; −2.38 and 4.55; 4.55 and 8.74; 8.74 and 12.80; 12.80 and 16.87; 16.87 and 20.96; or above 20.96; see Figure 2 and Table 3.

Table 3. derivation of autogenetic stages from index measure scores

| interpersonal adjective checklist - dominance		
score	**life satisfaction index**	
above	**below 13.66**	**above 13.66**
	(1) uninvolvement	
−24.76	(2) acceptance	(3) rebellion
−9.510	(4) necessity, obligation, responsibility	(5) escape
−2.380	(6) self-limitation, stinting	(7) choosing a limited environment
4.550	(8) manipulation	(9) opportunism
8.740	(10) single handed effort in partnership	(11) managing balance in partnership
12.80	(12) system functioning / determining aspects	(13) aware of system limitations
16.87	(14) independent / limiting dependent's activities	(15) independent / specifying one's activities regarding dependent
20.96		(16) independent and responsive
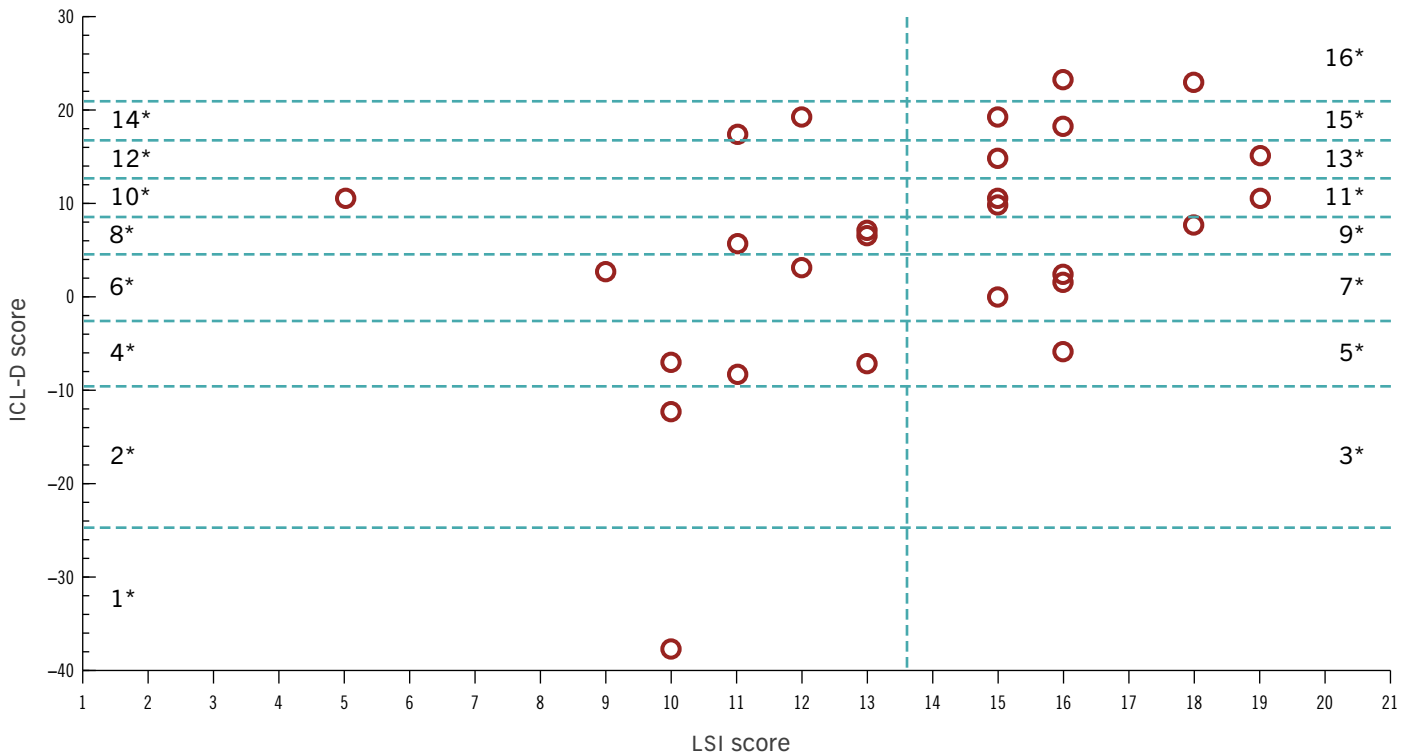
Figure 2. the index measure scores of autogenetic category of interview subjects
*Note: * autogenetic category numbers*

Confirming LSI and ICL-D as index measures of autogenesis

Interviews were not available for the Sage (s) self-help subjects. LSI and ICL-D scores were available for the Sage self-help subjects. It was hypothesized that the LSI and ICL-D scores for this sample of individuals would serve as index measures of autogenesis and that the autogenetic scores of the s Sample would diverge from those of the I sample in a theoretically predicted way. The autogenetic ratings of a sample (s) of individuals with a variety of overt interpersonal problems, resulting in their inability to successfully initiate or maintain themselves in vital personal relations, which they were addressing through membership in a self-help group, were computed utilizing the Index Measure of autogenesis. The Sage autogenetic ratings were compared with the autogenetic ratings of the original interview sample subjects who were selected to have scores distributed across the range of autogenetic groups. It was expected that the autogenetic ratings of the self-help group would differ from a sample in which the autogenetic ratings were evenly distributed, the assumption being that the self-help sample was less representative of a normal population distribution. It was expected that the majority of the self-help subjects would have autogenetic ratings at or below the ICL-D autogenetic index mean, $M = 9$, yielding autogenetic ratings of one to nine as opposed to ratings of ten to sixteen. A chi-square analysis found this to be the case, chi-square $(1, N = 42) = 6.4393$, $p < .01$; see Table 4).

Table 4. comparison of autogenetic categories of interview and self-help subjects above and at or below the autogenetic mean

autogenetic categories	subjects	
	interview	self-help
1–9	15	14
10–16	12	1

» DISCUSSION

Introduction

The purpose of this research was to confirm the validity of a model of interpersonal functioning named autogenesis, posited to account for the experience or lack thereof of a sense of life satisfaction in connection with interpersonal relations, presumed to evolve and to match to settings that inform interpersonal behavior. The transformations of these structures are posited to lead to broadened spheres of and personal depth in social participation, personal direction of one's activity, and personal and social responsibility.

The objectives of the research reported were three. 1) The research sought to establish the reliability of identifying autogenetic groups when examining individual life interviews. (2) It sought to affirm the combination of the LSI and ICL-D scores as an Index Measure of autogenetic categories. 3). It sought to confirm that a group characteristically identified with concepts reflecting some portion of the autogenetic sequence would primarily be composed of members whose individual autogenetic functioning fall within that portion of the sequence, demonstrating criterion validation of the autogenetic model.

Results

The Autogenetic Coding Manual developed to codify the varieties of autogenetic functioning that emerged from an analysis of I sample interviews proved upon refinement, highly reliable; enabling an independent rater to correctly identify with 93% agreement, $p < .001$, groups for which 11 of the interview subjects were identified by the manual developer.

This finding suggests that the autogenetic model has a theoretically consistent structure that can be perceived as having recognizable and distinguishable perspectives or categories that represent non-overlapping manners of expression, behaviors, and attitudes. Reliability of the manual, as representative of the autogenetic model, implies that the autogenetic model communicates distinctions in interpersonal perception and functioning.

If the autogenetic model informs interpersonal behaviors via a structure of interpersonal perception and is developmental or describes sequential elaboration, with less adaptive interpersonal stages preceding more adaptive ones, it was hypothesized that persons suffering from basic interpersonal failures involving establishing relations with others would function in the lower half of the autogenetic developmental sequence. The lower half of the A sequence falls below the partnership level, which is the first level above the midlevel of A stages. When the A stages of a sample (s) of individuals seeking help regarding their inability to initiate successfully interpersonally or maintain relationships were compared with a sample (i) of individuals evenly distributed across the autogenetic stages, 14 of the 15 s subjects were at stages below the partnership level, $p < .01$.

This finding is theoretically consistent with expectations and strongly suggests as predicted, the association of the categories with one another above and below the A sequence midpoint. The finding is limited in terms of verifying the interrelations of the stages within these groupings. This finding implies however, that the autogenetic sequence minimally represents groupings roughly defined in autogenetic terms as dependent and independent.

This finding also is consistent with the use of the combination of the LSI and ICL-D scores as an Index Measure of Autogenetic category.

The autogenetic model

The autogenetic model expresses in a symmetrical fashion the tensions between satisfaction and dissatisfaction and self and society associated with human experience. Freud (1963) in *Civilization and Its Discontents* suggests that the tension between self and society is permanent and irresolvable. An individual cannot pursue his own needs without coming into conflict with society. Theoreticians in the Freudian tradition have expressed the conflict inherent in this family of models in terms of life as equivalent to separation and individuation as opposed to union, fusion, dependency and death (Rank, 1929, 1945); autonomy as opposed to homonomy or surrender (Angyal, 1941, 1951, 1965); and the expression of agency as opposed to communion (Bakan, 1966, 1968, 1971).

Autogenesis presupposes that a human being is self oriented and whether other oriented as well from the outset or not, must determine his or her self realization in a social context. The context consists of opportunities and impediments to self-realization, which are vulnerable to change, and to which the individual must structure his interactive behavior in order to maintain his survival and attain self-realization. The end result of autogenetic development is an individual sense of initiative and self-realization that functions beneficially in a social context. The individual experiences the sense of meeting his needs and responding to others as non-conflictual.

The autogenetic model presently consists of 16 categories divided into nine levels. Within each level, with the exception of the first and last, there is one category representing satisfaction and one representing dissatisfaction. The levels represent the ever-expanding activity in and spheres of social involvement, to wit, (Level 1) Uninvolved, (Level 2) Acceptance/Rebellion, (Level 3) Obligation/Opportunism, (Level 4) Limitation, (Level 5) Manipulation/Opportunism, (Level 6) Partnerships, (Level 7) Systems, (Level 8) Independence, and (Level 9) Integration.

Progress through the levels of expanding personal activity and social spheres is marked by categories at each level expressing satisfaction and dissatisfaction. The categories of dissatisfaction entail an absence of a self-expressive posture and the categories of satisfaction entail the presence of a self-expressive posture. A comparison of the categories of dissatisfaction (uninvolved; accepting; sense of necessity; self-limitation; manipulation; silent effort in partnerships; working in a complex system of needs not fully understood; and establishing limits for dependents) with the categories of satisfaction (rebellion; escape; choosing a limited environment; opportunism; speaking out to attain balance in partnerships; functioning in a system, aware of its limitations; specifying one's activities vis-à-vis dependents; and independent and responsive) is indicative of the difference in functioning associated with the two types of categories.

Lack of expressiveness and action experienced as impacting on one are felt as unsatisfying. Self-expressive, autogenic action is felt as satisfying. It is these feelings of dissatisfaction paired with the absence of autogenic action and satisfaction paired with autogenic action that are theorized to propel the individual into contexts that embody consequences for vulnerabilities to, or perceived opportunities for, extending the course of autogenic action; that is for the continued restructuring of an individual's interpersonal behavioral structures.

An approach to examining the relationship between satisfaction and ego development was undertaken by Costa and McCrae (1983). Ego development stages, using Loevinger's (1976) sentence completion test, and measures of feeling of wellbeing were determined for each subject. No correlation was found between ego stage and feeling of wellbeing. However, as Bee (1987) points out "they did not check for the possibility of alternating higher and lower levels of wellbeing (p. 311)". The autogenetic model incorporates the concept of alternating higher and lower levels of wellbeing as integrally contributing to the process of interpersonal development.

The concept of alternations between dissatisfaction and satisfaction is reflected in the views of a number of theorists. The dialectical alternations in autogenetic theory are between self and other consideration and concerns, and activity and passivity, as well as satisfaction and dissatisfaction. Levinson (1978, 1980) describes the lives of adults as alternating between periods of upheaval and periods of stability in a non-dialectical way. James (1902) and Johnson (1983) see experience as more dialectical with recurring cycles of abandonment of existing organization in favor of reorganization. Kegan (1980) sees the experience of these alternations as painful. His view is that "Developmental theory gives us a way of thinking about such pain that does not pathologize it (p. 439)".

Bee (1987) recognizes development as spiral in form rather than linear. Her view is that "we continue to circle back to similar issues, but at higher and higher levels of decentering" (p.304). The spiral approximates a physical model of autogenesis, or interpersonally driven interpersonal development, with each ascending ring being wider than the last in recognition of the greater realms of interpersonal experience mastered at each succeeding level. This experience refers not only to the increasing inclusiveness of physical spheres in which interpersonal activities are undertaken, but to the increased breadth and depth of psychological experience.

The circumplex model from which the autogenetic model is derived

Kenneth Locke (2011) reports the interpersonal circumplex to have become the most popular model for conceptualizing, organizing, and assessing interpersonal dispositions or characteristics and describes what all examples of such models share. "The interpersonal circumplex is defined graphically by two orthogonal axes: a vertical axis (of status, dominance, power, control, or most broadly, *agency*) and a horizontal axis (of solidarity, friendliness, warmth, love, or most broadly, *communion*) (p.1). Horowitz (1996) previously described these models stemming from Leary's (1957) work similarly and noted "A number of factor analytic studies have confirmed this structure by showing that these two dimensions account for a large proportion of the variance in ratings of personality traits (p. 284)."

Agency and communion are central to a range of psychological theories, such as the theorizing by Blatt (2008) and self-determination theory (Bauer & McAdams, 2000). Agency and communion have a long history as constructs describing human interpersonal orientation and behavior as indicated previously. Cross-cultural research (Marcus and Kitayama, 1991, 1994) and the growth of multiculturalism in contemporary U. S. society have brought into the public domain the related cultural concepts of individualism and collectivism.

Locke (2011) notes that a variety of literatures support the centrality of agency and communion. Among these are the roles that hormones and neurotransmitters play in regulating communion through oxytocin (Bartz & Hollander, 2006), and agency through testosterone (Archer, 2006); supporting, in Locke's opinion, the idea that agency and communion are distinct tasks.

Although agency and communion may be viewed and studied as distinct tasks, there is research that suggests that agency and communion may become integrated across the lifespan. Mansfield and McAdams (1996) observe that generativity in adult lives can combine agency and communion. They found that highly generative adults in autobiographical expressions of agency and communion show greater levels of agency/communion integration.

Although the Leary ICL measure is scored separately for dominance (agency) and love (communion), only the dominance score in the Ethnic Study, from which the Interview Sample was drawn and on which the autogenetic model is based, correlates with life satisfaction. This correlation of life satisfaction with dominance as agency, suggests that agency is a core element of life satisfaction. Nonetheless, the autogenetic model of interpersonal development demonstrates that the integration of agency and communion is possible, in the forms of personal momentum and responsiveness to others, as realized in the most elaborated category of autogenesis. Agency and communion do not represent separate tasks; rather, according to the autogenetic model, a life long process of resolving, through multiple, context linked expressions of interpersonal functioning, the successful integration of the two.

Conclusion

The results of the research reported into the meaning of the Leary Interpersonal Adjective Checklist Scored for Dominance (ICL-D), in relation to life satisfaction, demonstrate the value of revisiting historically conducted research. In this instance, seminal measures such as the ICL-D and LSI can continue to contribute to an Index Measure of an interpersonal dynamic they were useful in identifying, autogenesis. The use of an Index Measure as an alternative or adjunct to the use of Interviews to research autogenesis can greatly facilitate research involving the autogenetic model. An Index Measure could also be useful in facilitating the identification of contemporary interpersonal circumplex measures as measures of autogenesis and perhaps as more decisive ones.

Also, in this instance of revisiting historically conducted research, the identification of the model of autogenesis proves useful for addressing contemporary questions regarding the integration of the interpersonal constructs of agency and communion; and for opening lines of research to investigate these integrative processes and identify opportunities for their fruitful application. ■

REFERENCES

Acton, G. S. (2002). Interpersonal measures show circumplex structure based on new psychometric criteria. *Journal of Personality Assessment*, 79, 446–471.

Adams, D. (1969). Analysis of a L.S.I. *Journal of Gerontology.* 470–474.

Angyal, A. (1941). *Foundations for a science of personality*. New York: Commonwealth Fund.

Angyal, A. (1951). A theoretical model for personality studies. *Journal of Personality*, 20, 131–142.

Archer, J. (2006). Testosterone and human aggression: an evaluation of the challenge hypothesis. *Neuroscience and Biobehavioral Reviews, 30,* 319 – 345.

Bakan, D. (1966). *The duality of human experience.* Chicago, IL: Rand McNally.

Bakan, D. (1968). *Disease, pain and sacrifice.* Chicago, IL: University of Chicago Press.

Bakan, D. (1971). *Slaughter of the innocents.* San Francisco, CA: Josey-Bass.

Bee, H. E. (1987). *The journey of adulthood.* New York, NY: Macmillan.

Bauer, J.J., & McAdams, D.P. (2000). Competence, relatedness, and autonomy in life stories. *Psychological Inquiry, 11,* 276–279.

Blatt, S. J. (2008). *Polarities of experience: Relatedness and self-definition in personality development, psychopathology, and the therapeutic process*. Washington, DC: American Psychological Association. doi:10.1037/11749–000

Craig, R. J. (2005). Assessing personality and mood with adjective check list methodology: A review. International *Journal of Testing, 5*(3), 177–196.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience.* New York, NY: Harper and Row.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Costa, P.T., Jr., McCrea, R.R., & Norris, A.H. (1983). Recent longitudinal research on personality and aging. In K.W. Schaie (Ed.), *Longitudinal studies of adult psychological development.* New York, NY: Guilford Press.

Bartz, J. A., & Hollander, E. (2006). The neuroscience of affiliation: Forging links between basic and clinical research on neuropeptides and social behavior. *Hormones and Behavior, 50,* 518–528.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment, 49,* 71–75.

Diener, E., Inglehart, R., & Tay. L. (2013). Theory and validity of life satisfaction scales. *Social Indicators Research, 112*(3), 497–527. Netherlands: Springer. doi: 10.1007/s11205–012–0076-y

Diener, E., Oishi, S. & Lucas, R. E. (2009). Subjective well-being: The science of happiness and life satisfaction. In C.R. Snyder & S.J. Lopez (Eds.). *Oxford Handbook of Positive Psychology* (2nd Ed.). New York, NY: Oxford University Press.

Fowler, J.W. (1981). *Stages of faith.* New York, NY: Harper & Row.

Freud, S. (1963). Civilization and its discontents. London: Hogarth Press.

Hattie, J. (2009). *Visible learning.* London: Routledge.

Hemphill, J.F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist, 58*(1), 78–80. doi: 10.1037/0003–066X.58.1.78

Horowitz, L. M. (1996). The study of interpersonal problems: A Leary legacy. *Journal of Personality Assessment, 66*(2), 283–300.

James, W. (1902). The varieties of religious experience. New York: NY: Mentor.

Johnson, D. (1983). Spirals of growth. Wheaton, IL: Theosophical Publishing House.

Kegan, R. (1980). There the dance is: Religious dimensions of developmental theory. In J.W. Fowler & A. Vergote (Eds.), *Toward moral and religious maturity.* Morristown, NJ: Silver Burdett.

Kohlberg, L. (1964). *Collected papers on moral development and moral education.* Cambridge, MA: Center for Moral Education.

LaForge, R., & Sugzek, R. F. (1955). The interpersonal dimension of personality: III an interpersonal checklist. *Journal of Personality*, 24(1), 94–113.

Leary, T. (1996). Commentary. *Journal of Personality Assessment, 66*(2), 301–307.

Leary, T. (1957). *Interpersonal diagnosis of personality.* New York: Roland Press.

Lenhard, W., & Lenhard, A. (2014). *Calculation of Effect Sizes.* Retrieved from http://www.psychometrica.de/effect-size.html

Levinson, D.J. (1978). The seasons of a man's live. New York, NY: Knopf.

Levinson, D.J. (1980). Toward a conception of the adult life course. In N.J. Smelser & E.H. Erikson (Eds.), *Times of work and love in adulthood*. Cambridge, MA: Harvard University Press.

Liberman, M., & Cohler, B. (1975). *Constructing measures for older people*. (Final Report A.O.A. Grant No. 93-p57425/5). Chicago, IL: Authors.

Locke, K.D. (2011). Circumplex Measures of Interpersonal Constructs. In L. M. Horowitz & S.

Strack (Eds.), *Handbook of Interpersonal Psychology.* Hoboken, NJ: John Wiley and Sons Inc.

Loevinger, J. (1966). The meaning and measurement of ego development. *American Psychologist, 21*, 195–206.

Loevinger, J (1976). *Ego Development.* San Francisco, CA: Josey-Bass.

Loevinger, J (1984). On the self and predicting behavior. In R.A. Zucker, J. Aronoff, & A.I. Rabin (Eds.) *Personality and prediction of behavior.* New York, NY: Academic Press.

Mackinnon, S.P., Sherry, S.B., & Pratt, M.W. (2013). The relationship between perfectionism, agency, and communion: A longitudinal mixed methods analysis. *Journal of Research in*

Personality 47, 263–271. doi:10.1016/j.jrp.2013.02.007

Mansfield, E.D., & McAdams, D.P. (1996). Generativity and themes of agency and communion in adult autobiography. *Personality and Social Psychology Bulletin*, 22(7), 721–731.

Markus, H.R., & Kitayama, S. (1994). A collective fear of the collective: Implications for selves and theories of selves. *Personality and Social Psychology Bulletin, 20,* 568–579.

Markus, H., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98,* 224–253.

Neugarten, B. Havighurst, R., & Tobin, S. (1961). The measurement of life satisfaction. *Journal of Gerontology.* 1(2), 134–143.

Nordmann, N. (1991). *An autogenetic approach to interpersonal development* (Unpublished doctoral dissertation). University of Chicago, Chicago, IL.

OECD Better Life Index. *Life Satisfaction.* The Organization for Cooperation and Development. Retrieved from http://www.oecdbetterlifeindex.org/topics/life-satisfaction

Pew Research Center (2014). *People in emerging markets catch up to advanced economies in life satisfaction.* Pew Research Global Attitudes Project. Retrieved from www.pewglobal.org/2014/10/30/people-in-emerging-markets-catch-up-to-advanced-economies-in-life-satisfaction/

Rank, O. (1929). *The trauma of birth.* New York, NY: Harcourt, Brace.

Rank, O. (1945). *Will, therapy and truth and reality*. New York, NY: Knopf.

Selman, R.L. (1980). *The growth of interpersonal understanding: Developmental and clinical analysis.* New York, NY: Academic Press.

Strack, S. (1996). Introduction to the special series—interpersonal theory and the interpersonal circumplex: Timothy Leary's Legacy. *Journal* of *Personality Assessment, 66*(2), 212–216.

Sullivan, H.S. (1953). *The interpersonal theory of psychiatry*. New York: W.W Norton & Co.

Wiggins, J.S. (1995) *Interpersonal Adjective Scales: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Wiggins, J.S., Trapnell, P., & Phillips, N. (1988). Psychometric and geometric characteristics of the Revised Interpersonal Adjective Scales (IAS-R). *Multivariate Behavioral Research, 23*, 517–530.