Special issue on tests and measures of positive adult development

BEHAVIORAL DEVELOPMENT BULLETIN

BEHAVIORAL DEVELOPMENT BULLETIN

A Behavior Analysis Online Journal & An American Psychological Association Publication

VOLUME **19**   NUMBER **3**   SEPTEMBER **2014**

» 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

» 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 six 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. Precision teaching and celeration

6. 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

» SECTION 1: THEORY AND MODELS

» SECTION 2: INSTRUMENTS AND APPLICATION

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.

The aim of the present study was to describe, test and validate a method for disclosing significant response patterns from questionnaire data, and for classifying individual response profiles into a sequence of significant patterns. The method is based on pattern recognition statistics and probability calculations. The results from the population tested show that the method can disclose characteristic profiles of different value systems, and that these systems can be arranged in a hierarchical order similar to the conventional levels of ego development. It is suggested that this method is applicable to any multiple choice-questionnaire containing a number of items where the response alternatives represent a sequential order, for example, of different levels of development within a psychological domain. The method might be a valuable tool for acquiring information on the distribution of different levels of adult development in large populations, such as in communities and large organizations.

Preface

Michael Lamport Commons, Editor
Harvard Medical School

This Special issue of the *Journal of Adult Development* will focus on the measurement of positive changes during adulthood. Adult development has multidimensional threads (Commons, 1999). These threads are sequence of events that run throughout the whole course of adult life. Most of previous work on measuring change in adulthood examines 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. 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. Most 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 Werner.

The first section of this issue includes articles that a paper in which the MHC stages have been revised. It also includes a paper that presents the formal definitions and axioms of the Model of Hierarchical Complexity (MHC). This model is a breakthrough in the field of development as it is a general stage model that can be used to study development in any domain, including adult development. 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 second section of this issue is a compilation of articles on 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 moral judgment, caring, physics, social, corruption and so on.

In the first section of the issue, the opening article is the first revision of the order and the corresponding stage sequence of the Model of Hierarchical Complexity since it was created in 1982. Although it may be difficult to remember the new numbers, the names have all stayed the same. The original Sensory or Motor Order 1 has now been separated into Automatic Order 1 and Sensory or Motor Order 2. This revision was made because of the discovery that classical conditioning is more hierarchically complex than sensitization and habituation. The second paper formalizes the Model of Hierarchical Complexity by presenting its formal definition and axioms. The next paper discusses four different accounts of stage transition, each delineating how to obtain data on stage transition for each method. It also discusses how those methods might be interrelated. The forth paper is a commentary in which calls for a need to combine this new model of Boom (2012) and MHC to generate a complete model of stage development. The fifth paper constructs a very useful correspondence table comparing the stages of five developmental stage models, which cover the lifespan of an individual. In addition, this paper provides the translation of Colby and Kohlberg's (1987a, 1987b) 13 point scale into the other developmental scales that has not been done before. The five scales included in this table are MHC (Commons, Richards & Armon, 1984), Fischer and Bidell's (1998) Stages of Cognitive Development, Piaget and Inhelder's (1969) Stages of Cognitive Development, Colby and Kohlberg's (1987a, 1987b) 9 Point Scale of Moral Judgment and Colby and Kohlberg's (1987a, 1987b) 13 Point Scale of Moral Judgment. The sixth paper proposes that integrating MHC and the nonlinear model of the fractal transitional orders of hierarchical complexity can help develop a unified mathematical theory of behavioral development. The final paper of this section is about the transition to meta-cross-paradigmatic stage, stage 16, of the Orders of Hierarchical Complexity.

The opening article of the second section is by Golino, Mauro, Commons and Miller on construction and validation of the Inductive Reasoning Developmental Stage test (IRDT), an instrument based on MHC, used to identify developmental stages of individuals. The next paper provides supporting evidence for Defining Issues Test (DIT) as a measure of moral judgment development. The third paper shows how the Model of Hierarchical Complexity can be applied to physics. It shows implications of understanding the history of wave theory in physics from an adult development perspective including the MHC. The fourth paper reviews literature on how literature on corruption does not offer a multi-faceted and integrative understanding of the phenomenon. The fifth paper goes on to apply the MHC to provide a synergetic, interdisciplinary framework for analyzing and evaluating corruption and the discourse dealing with it. The fourth and fifth papers suggest implications of viewing corruption from the perspective of adult development and Model of Hierarchical Complexity. The sixth paper presents a recently developed instrument of care-based moral development: The Ethic of Care Interview (ECI) and provides evidence for the importance of care-oriented morality for human growth, especially identity and personality development. The seventh paper uses MHC as a framework to study individual's stages of moral understanding. Finally, this issue concludes with the last paper on testing and validating the method for disclosing signal response patterns from questionnaire data and for classifying individual response profiles into a sequence of significant patterns. This method makes use of the MHC.

Lastly, I would like to thank my managing editors, Eva Yujia Li and Charu Tara Tuladhar, for their help in putting this issue together.

**BEHAVIORAL
DEVELOPMENT
BULLETIN**

Introducing a new stage for the model of hierarchical complexity: A new stage for reflex conditioning

Michael Lamport Commons[1] and Thomas Ruofei Jiang[2]

1 Harvard Medical School
2 Harvard University

The model of hierarchical complexity (MHC) is known to have 16 orders so far. However, applying the model to explain the development of operant conditioning (*original order 2*) from respondent conditioning (*original order 1*) in non-human animals has led to the recent discovery of a new stage.

Actions that make up respondent conditioning are more hierarchically complex than habituation, sensitization, and other simple actions or behavioral tendencies that were also included in *original order 1*. Thus, the *original order 1* has now been separated into the new *automatic order 1* and the new *sensory or motor order 2*. All the orders above the *original order 1* also had their numbers incremented by one. Thus, there are now 17 orders of hierarchical complexity. This paper describes this new sequence of orders at the lowest end of the model.

KEYWORDS: revised, orders, model of hierarchical complexity, stages

The addition of a New Order 2 between the Revised Order 1 and the Original Order 2 has a curious history. Kurt Fischer (personal communication, June, 1984) suggested that there has always been a possibility that there are more stages in infancy than the ones foreseen in the Model of Hierarchical Complexity. There would actually be a couple of reasons for this. One is that Piaget had originally proposed six stages during the first 18 months of life. The other is that original stage schemes were not conceived of with non-human animals in mind. In our work, we have been applying the Model of Hierarchical Complexity to explain the development of operant conditioning (Original Order 2) from respondent conditioning (Original Order 1) in non-human animals. This led to the discovery that actions making up respondent conditioning were more hierarchically complex than habituation, sensitization, and other simple actions or behavioral tendencies also included in Original Order 1. This realization led to the insertion of a New Sensory or Motor Order 2 that follows the renamed and revised Automatic Order 1 to account for the evolution and development of respondent conditioning. The revised Automatic Order 1 goes between Original Calculatory Order 0 and the new Sensory or Motor Order 2. The Original Order 1 then became Sensory or Motor Order 2, and all the orders above that also had their numbers incremented by one. This paper is about this new sequence of orders at the lowest end of the Model.

The model of hierarchical complexity and its axioms

The Model of Hierarchical Complexity (MHC) is a mathematical model that sets forth a measurement system by which actions are put into a hierarchical order. The model assesses a general, unidimensional developmental measure of difficulty across domains. The Model of Hierarchical Complexity suggests that one of the major ways in which sequences of tasks are arranged is in terms of their complexity (or difficulty). The complexity of a task is operationalized in terms of its Order of Hierarchical Complexity (OHC). The measurement system of the model is composed of axioms. Axioms are rules that are followed to determine how the model orders actions to form a hierarchy. There are five axioms:

Axiom 1 (Well-ordered). If one action is less complex than another action, then the assignment function, which gives a numerical order of hierarchical complexity to an action, must preserve the action's order in the hierarchy. In non-mathematical terms: That is, simpler actions are lower in the order than more complex actions.

Axiom 2 (Transitivity). If action a is more hierarchically complex than action b, and action b is more hierarchically complex than action c, then action a is more hierarchically complex than action c.

Axiom 3 (Chain rule). When actions a and b are chained together in some order, and the order in which they are executed is not influential to accomplishing a task, the order of hierarchical complexity of $(a \circ b)$ equals that of the highest subaction. In

non-mathematical terms: That is, when two actions, a and b, are organized in some way, but the actions can be completed in any order, then the overall hierarchical complexity of the two chained actions is only as high as the most hierarchically complex action in the chain.

Axiom 4 (Coordination rule). The organization of the ordering of action rules is non-arbitrary. In non-mathematical terms: When two actions, a and b, are organized, that organization has to be non-arbitrary.

Axiom 5 (Equal spacing[1]). The a priori difficulty of a task action changes by 1 for each change in the Order of Hierarchical Complexity, irrespective of what adjacent Orders of Hierarchical Complexities one is comparing. In other words, there is equal spacing between each order

Original order 1 violates axiom 1

Originally, Sensory or Motor Order 1 was defined as an order in which organisms coordinate one action or operation with one stimulus. They engage in a single action at a time and the action is not coordinated with other actions, but with a stimulus. Both the detection of stimuli and the production of responses are somewhat flexible, but the relationship between them is not. This order was described as including actions such as reflexes, sensitization, habituation, tropisms and last but more troubling, respondent conditioning.

The problem is that respondent conditioning cannot belong to Original Sensory of Motor Order 1 or Original Circular Sensory-motor Order 2. Although respondent conditioning was previously categorized as Original Order 1, it is more hierarchically complex than the Original Order 1. Respondent conditioning requires that a neutral stimulus NS be changed into a conditioned stimulus CS. This involves the procedural pairing of a presently neutral stimulus NS that only elicits attention with an unconditioned stimulus UCS that elicits an unconditioned response UR. According to Axiom 1 of the MHC, actions at the next higher Order of Hierarchical Complexity are defined in terms of two or more actions from the adjacent next lower order. In this case, actions from the adjacent next lower orders include: 1) attentional response to the neutral stimulus NS and 2) unconditioned response UR to unconditioned stimuli UCS. Once the neutral stimuli NS and unconditioned stimuli UCS are procedurally paired (ordered), the neutral stimulus becomes the conditioned stimulus CS that elicits a conditioned response CR. Thus, the conditioned response CR is more hierarchically complex than either the attentional response to a neutral stimulus or the unconditioned response UR. Hence, it is vital to separate Original Order 1 into two different orders, in which, the lower order includes actions such as unconditionable

reflexes, sensitization, habituation and tropisms and the higher order includes respondent conditioning.

Likewise, it will be argued that operant conditioning belongs to new Order 3. This is because operant conditioning in our account is built out of three instances of respondent conditioning

» THE NEW ORDERS

The Original Sensory or Motor Order 1 has now been divided into the new Automatic Order 1 and the new Sensory or Motor Order 2. Originally, there were 16 Orders of Hierarchical Complexity. With the insertion of the new Order 2, there are now 17 Orders of Hierarchical Complexity as shown in Table 1.

The need for a new Order 2 was discovered while reviewing observational and experimental literature on animal behavior in order to determine the behavioral developmental stages at which those animals performed. It was found that single celled organisms did not classically condition. When the literature on classical conditioning on single celled organisms was reviewed, it was found that the behaviors exhibited were habituation and sensitization. No neutral stimulus (NS) was conditioned. This suggested that habituation and sensitization could not be in the same order as classical conditioning.

This section presents the updated first four orders of hierarchical complexity. They are illustrated in Table 2 using examples drawn from observational and experimental literature on animal behavior. The table is followed by elaborate descriptions of each order.

Calculatory order 0

Order 0 includes pre-programmed behaviors that are very specifically elicited by exact computations. The forms of responses do not show variation and the responses to a "stimulus" show no generalization. There is no gradated response. The behavior is not elicited by any form of intelligent acting organism or thing. For example, a computer program behaves at stage 0. In a computer program, codes are initially provided by human programmers. Programmers perform at a stage that is incredibly higher than the computer programs do. What the program does is fixed and cannot be changed without a programmer. Of course there is programmed machine learning, but even small random changes in the stimulus or response are not possible. Similarly, in biology, the behavior of deoxyribonucleic acid (DNA) is at Order 0 because it performs a biological "calculation or programming" that happens almost the same way every time. In other words, for nucleotides, C always "bonds" with G and A always bonds with T, similarly to binary coding. We are not concerned with the biochemistry, but just the genetic code and the behavior of the nucleotide bases.

Automatic order 1

For most of evolutionary time, there were only single-celled organisms. From our review it makes sense to assume that single-celled or-

Table 1. The updated orders of hierarchical complexity

Order number	Order name
0	Calculatory
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

1. Optional

ganisms in the evolutionary past also only had "hard wired" responses including taxis, tropisms and phagocytosis and the like (Commons & White, 2006/2009).

The Automatic Order 1 is a very slightly modified version of Original Sensory or Motor Order 1. The only change was the removal of respondent conditioning. The criterion for Automatic Order 1 is that the organism engages in a single action at a time and the action is "hard wired" into the organism. Single celled organisms respond to a single environmental stimulus. Responses to naturalistic events occur because these hard wired actions are tuned to certain relatively specific stimuli. The environmental stimulus S that leads to the behavior is not paired with any other stimulus. The single action is an innate biological action to a specific environmental stimulus. Examples of the environmental stimulus S could be a chemical emitted by possible food, light, heat, or electricity. The actions are built into the organism. Examples of such built in or automatic actions include taxis, tropisms, phagocytosis and unconditionable reflexes (Commons & White, 2006/2009). Obviously, single celled animals do not have nervous systems.

Here, *conditionable* and *unconditionable* reflexes are distinguished. Uncontionable reflexes are an Order 1 behavior. Reflex, is nearly an instantaneous movement in response to a stimulus (Purves, 2004). In an unconditionable reflex, the stimulus and the response are coordinated, and the coordination is totally automatic. Reflexes that are not classically conditioned are Automatic Order 1 responses. They will be referred to as unconditionable reflexes. Also, the term reflex is used here, as opposed to tropism or taxis because the term reflex is traditionally used for fast responses that do not have long durations. Reflexes that are classically conditioned will be referred to as conditionable reflexes, which are Sensory or Motor Order 2 response.

Simple learning such as habituation and sensitization are also Automatic Order 1 actions that have been shown to occur. This learning is distinct from later forms in that while changes in behavior do occur, they only occur in response to changes in those specific stimuli to which those behaviors generally respond. These are two forms of non-associative learning. These are behavioral processes that may have evolved to deal with stimuli that occur iteratively in the environment (Eisenstein, Eisenstein & Smith, 2001). Habituation is a decrease in magnitude of a response to an iterative stimulus. On the other hand, sensitization is an increase in magnitude of a response to an iterative stimulus. These forms of learning are distinct from later forms of classical conditioning, sometimes called associative learning. Single celled organisms at Order 1 have limited sensors and effectors. There are no uncontroversial reports of such organisms responding in actions above Order 1.

Some examples of order 1 animals. Order 1 actions will be illustrated using examples from studies on paramecia, protozoan; *Vorticella convallaria*, and protozoan *Spirostomum*.

Example 1. This is an example of unconditionable reflex and habituation as an Automatic Order 1 behavior in protozoan, *Vorticella convallaria* by Patterson (1973).

Stimulus 1 (s_1). Electric stimulation of different intensities administered every 10 seconds for 5 minutes.

Response 1 (R_1). Response to s_1, was contraction of the body and stalk.

s_1 eliciting R_1 is an example of unconditionable reflex which is an Automatic Order 1 behavior.

Stimulus 2 (s_2). Mechanical stimulus administered by dropping different weights on the microscope stage every 10 seconds for 5 minutes.

Response 1 (R_1). Response to s_2 was contraction of the body and stalk.

s_2 eliciting R_1 is also example of unconditionable reflex which is an Automatic Order 1 behavior.

Stimulus 3 (s_3). Mechanical stimulus was administered by modifying the media of the organism.

Response 1 (R_1). Response to s_3 was contraction of the body and stalk.

s_3 eliciting R_1 is also example of unconditionable reflex which is an Automatic Order 1 behavior. Habituation occurred with administration of all the three stimuli. The longer the organisms were exposed to the stimuli, the longer became the periods in which the organism were non-responsive.

Table 2. Revised description of the first five orders of hierarchical complexity

Order name	Order #	Task	How it is done	Who does it
Calculatory	0	Follow computer program; dna; calculate; store information	Manipulate 0, 1; four nucleotide bases	Human made program;
Automatic	1	Reflexes, sensitization, habituation, tropisms	Engages in one action at a time. Cellular activities: sensing, effecting	Single celled organisms
Sensory or motor	2	Reflexes and respondent conditioning	Procedurally pair an unconditioned stimulus (ucs) that elicits an unconditioned response (ur) with a salient neutral stimulus (ns)	Animals with very simple nervous systems, slugs, leeches, some mollusks
Circular sensory motor	3	Operant conditioning	Coordinate three steps of respondent conditioning	Animals with a nervous system: some worms, insects
Sensory-motor	4	Learn concepts	Coordinate two or more operant	Mammals, birds, reptiles

Example 2. Paramecia are Automatic Order 1 animals. This is shown by their failure to classically (Mingee, 2013) and operantly condition (Mingee & Armus, 2009). They show behaviors of sensitization.

Stimulus 1 (s_1). One of the stimuli used in the study by Mingee (2013) was level of illumination.

Response (R_1). Response to s_1, level of illumination, was moving away from light (in most paramecia with the exception of Paramecia *bursaria*).

s_1 eliciting R_1 is an example of taxis which is an Automatic Order1 behavior.

Stimulus 2 (s_2). The other stimulus used was shock in the cathode side of the trough.

Response (R_2). Response to s_2 was swimming to the non-cathode side (i.e., moving away from the shock).

s_2 eliciting R_2 is also example of taxis which is an Automatic Order1 behavior.

When s_1 and s_2 were paired to investigate whether s_1 would elicit the same response as s_2 after the pairing (i.e., checking for presence of classical conditioning), it was found that s_1 no longer elicited R_2 after 1 minute of the first testing trial. Thus, pairing of the two stimuli was unsuccessful and classical conditioning did not occur suggesting that paramecia behave at Automatic Order 1.

Example 3. This is an example of unconditionable reflex, habituation and sensitization as an Automatic Order 1 behavior in protozoan *Spirostomum ambigum* in the study done by Hamilton, Thompson and Eisenstein (1974).

Stimulus 1 (s_1). Vibratory stimulus was administered for 10 minutes repetitively (0.1 HZ)

Response 1 (R_1). Response to s_1, vibration stimulus, was contractions, rapid shortening of the organism to about one-half of its resting length.

s_1 eliciting R_1 is an example of unconditionable reflex which is an Automatic Order 1 behavior .

The organisms that were initially less reactive (contracted less frequently) showed sensitization whereas, the organism that were initially more reactive habituated. These results were replicated by Eisenstein, Brunder and Blair (1982).

Organisms behaving at Order 1 would be insensitive to outcomes except in an evolutionary sense. That is, consequences may be selected for in an evolutionary sense if the single response leads to survival and reproduction.

Sensory or motor order 2

At Sensory or Motor Order 2, organisms coordinate two stimulus response pairs from the lower Automatic Order 1. An example of this is respondent conditioning. The criterion for classifying something as Sensory or Motor Order 2 is that the pairing of

stimuli leads to conditioning (Commons, Miller, Commons-Miller & Chen, 2012). Unlike at Order 1, the responses begin to be more flexibly associated with stimuli with which they have been paired. Either the detection of stimuli or the production of responses is somewhat flexible.

For organisms performing at Sensory or Motor Order 2, the important forms of behavior for the account being presented here are reflexes and the most complex process is respondent conditioning.

A *reflex* procedurally links stimulus to response (Pavlov, 1927). Reflexes can be mediated by a reflex arc only a few neurons long (Palkovits & Záborszky, 1977). In a reflex, the stimulus and the response are coordinated, but the coordination is automatic. For example, when water moves, mollusks open their shells reflexively (Palkovits & Záborszky, 1977). If something touches their membrane, the shells close. There is very little variability in these responses.

For a respondent conditioning procedure, a Sensory or Motor Order 2 task action is the "pairing" of two eliciting stimuli: an Environmental Stimulus (s) and an Unconditioned Stimulus (UCS). A salient UCS and S already exist before the pairing and the endogenously salient UCS automatically elicits the unconditioned response (UCR). After a sufficient number of occurrences, such pairings transform the neutral stimulus (s) into a conditioned stimulus (CS). The CS becomes more salient by having acquired most of its saliency from being paired with the endogenously salient UCS (Lawrence, Klein & LoLordo, 2009). This CS then elicits the conditioned response (CR), which is a variation of the unconditioned response (UR) (Pavlov, 1927). In respondent conditioning, there is the organization of stimulus elicited actions by organizing the stimuli.

The transfer of salience is at Sensory or Motor Order 2 of Hierarchical Complexity because: *a*) two stimuli are arbitrarily paired either by accident or by an experimenter, *b*) the organism's behavior does not directly cause the reinforcing stimuli in this situation as it does in operant conditioning, and *c*) the organism does not temporally or in some other way organize or coordinate more than one action in order to more adequately accomplish this task. Therefore, this pairing of the s and US does not constitute an increase in the hierarchical complexity of the task that must be solved. Using the example above, each of the arbitrary pairings of two salient stimuli that make up the three procedural steps meets the criteria for Sensory or Motor Order 2 in the MHC

To perform Sensory or Motor Order 2 task actions, organisms have to have networks of neurons to organize the conditioning of reflexes. As it is likely that the existence of neurons dates to slightly before the Cambrian period, we speculate that organisms, which at a minimum respondently conditioned, developed not much before or during the Cambrian explosion. This speculation is based on the fact that prior to the Cambrian explosion, most organisms were simple, composed of individual cells occasionally organized into colonies (Butterfield, 2001). Then, in the Cambrian explosion, there was the relatively rapid appearance of most major animal phyla. Among the animals that evolved during that period were the chordates, animals with a dorsal nerve cord; hard-bodied brachiopods, which resembled clams; and arthropods, ancestors of spiders, insects and crustaceans.

Some examples of sensory or motor order 2 actions. Order two actions will be illustrated using examples from three studies.

Finding current animals that respondently condition but do not operantly condition is a difficult one. That is partly because many people who have been studying invertebrates in particular, who are candidates for being this kind of animal, have been primarily interested in doing neuronal studies of these relatively simple animals as they are undergoing classical conditioning (Abramson, 1994). For most of the instances of classical conditioning that we have come across, we just do not know whether operant conditioning of that organism has even been attempted. In most cases, no published reports have been found. That does not of course mean that attempts have not been made.

Example 1. The first example comes from the study done by Henderson and Strong (1972) on *Macrobdella ditetra* (leech). In the study, they successfully classically conditioned leeches.

Neutral stimulus (NS). The neutral stimulus NS used in this study was light from light bulb.

Neutral response (NR). Neutral response to NS, light, was cephalic turning response. This is a natural response to light.

Unconditioned stimulus (UCS). The unconditioned stimulus UCS used in this study was shock.

Unconditioned response (UR). The unconditioned response UR was the anteroposterior contraction after the presentation of UCS. This is the natural response to shock.

Neutral stimulus and unconditioned stimulus pairing. The neutral stimulus (NS), light, was paired with the unconditioned stimulus (UCS), shock. The NS was presented for 3 seconds and then the UCS was presented for 0.1 second during the last 0.1 second of the NS.

Conditioned stimulus (CS). After the NS and UCS pairing, light became the conditioned stimulus.

Conditioned response (CR). After the light became a conditioned stimulus, it elicited the same response as the UR did which was anteroposterior contraction during CS, but before UCS. Thus, anteroposterior contraction became the CR and the light no longer elicited the NR.

In this example, light (NS) eliciting cephalic turning response (NR) in leeches is one automatic order 1 action. The second automatic order 1 action was the shock (UCS) eliciting anteroposterior contraction (UR). These two order 1 actions are coordinated (paired) to form the Sensory or Motor order 2 action which is light (CS) eliciting anteroposterior contraction (CR).

Example 2. The second example planarian, dugesia dorotocephalau, were classically conditioned by Thompson and McConnell (1955).

Neutral stimulus (NS). The neutral stimulus NS used in this study was light from light bulb.

Neutral response (NR). Neutral response NR to, light NS, in the control animals was low (10–30%) rate of turn responses, and a very low (<5%) contraction rate.

Unconditioned stimulus (UCS). The unconditioned stimulus UCS used in this study was shock.

Unconditioned response (UR). The unconditioned responses UR were a sharp turning of the cephalic region to one side or the other, and a longitudinal contraction of the entire body.

Neutral stimulus and unconditioned stimulus pairing. The neutral stimulus (NS), light, was paired with the unconditioned stimulus (UCS), shock. The NS of light was presented for 3 seconds and then the UCS of shock was presented for 1 second during the last 1 second of the NS.

Conditioned response (CR). After the light became a conditioned stimulus CS, it elicited the same responses as the UR did which were a sharp turning of the cephalic region to one side or the other, and a longitudinal contraction of the entire body.

In this example, light (NS) rarely eliciting a turning or contracting response (NR) in planarian is one automatic order 1 action. The second automatic order 1 action was the shock (UCS) eliciting a higher probability turning or contracting response (UR). These two order 1 actions are coordinated (paired) to form the Sensory or Motor order 2 action which is light (CS) eliciting a higher probability turning or contracting response (CR).

Example 3. The third example comes from the study done by Mpitsos and Davis (1973) on marine gastropod Pleurobranchaea (sea slugs). In the study, they successfully classically conditioned sea slugs.

Neutral stimulus (NS). The neutral stimulus NS used in this study was tactile stimulation of the oral veil using a sterile glass probe.

Neutral response (NR). Neutral response to NS, tactile stimulation of the oral veil, was withdrawal and bite-strike response.

Unconditioned stimulus (UCS). The unconditioned stimulus UCS used in this study was food chemicals (Homogenized squid).

Unconditioned response (UR). The unconditioned response UR was feeding behavior after the presentation of UCS.

Neutral stimulus and unconditioned stimulus pairing. The neutral stimulus (NS) was paired with the unconditioned stimulus (UCS), food chemicals. The NS (sterile glass probe for tactile stimulation) was coated with the food chemicals, UCS, and the oral veil was stroked for 10 seconds.

Conditioned stimulus (CS). After the NS and UCS pairing, tactile stimulation of the oral veil became the conditioned stimulus.

Conditioned response (CR). After the tactile stimulation of the oral veil became a conditioned stimulus, it elicited the same response as the UR did which was feeding behavior during CS, but before UCS. Thus, tactile stimulation of the oral veil became the conditioned response and the tactile stimulation of the oral veil no longer elicited the NR.

Circular sensory motor order 3

At Circular Sensory Motor Order 3, organisms coordinate two or more actions from Sensory or Motor Order 2. The most important case is that of Operant Conditioning. Operant Conditioning may be accounted for by the three steps of procedural respondent conditioning. Organisms that solve Circular Sensory Motor Order 3 tasks are multi-celled with some sort of more complex nervous system than what is seen in Sensory or Motor Order 2 animals. This section presents an argument that operant conditioning is Circular Sensory Order 3 action. Operant conditioning results from the coordination or organization of three respondent conditioning steps. These steps are: step 1, "What to do"; step 2, "When to do it"; and step 3, "Why to do it".

In Step 1, there is an assumed *representation of behavior* that elicits entering the hole. That *representation of behavior* becomes salient by being paired with the sucrose reinforcement, ucs/s^{R+}.

In Step 2, we understand that Sokolowski Disma and Abramson (2010) indirectly showed that the now salient *representation of behavior*, which elicits the operant behavior, R, is paired with the environmental stimulus, s (the turning on of the LED lights around the hole). Here the operant behavior R, is entering the hole to get to the reinforcement.

In Step 3, the environmental s (the visible hole with LED lights around it) is paired with the sucrose reinforcement, ucs/s^{R+} making the s more salient and valuable. This pairing acts to produce an incentive. The environmental s takes on the elective properties of ucs/s^{R+}.

Specifically the three steps of respondent conditioning are from Order 2 as required by the axioms of the MHC. At Order 2, the pairing at each step of procedural respondent conditioning occurs independently of the other respondent conditioning steps. Those steps are not coordinated at that order.

Order 3 — Examples

What follows, are some examples of operant conditioning in insects. Insects and some related animals were chosen to show how Order 3 Operant Conditioning may be accounted for by the three steps of procedural respondent conditioning.

Some examples of order 3 actions. Order three actions will be illustrated using examples from three studies. Order three actions will be shown to coordinate three Sensory or Motor Order 2 actions.

Example 1. Sokolowski et. al (2010), showed that blowfly (Protophormia terrae novae) behavior can be operantly conditioned. In this example, Steps 1, 2, and 3 are illustrated by what happens when blowfly behavior is operantly conditioned.

Individual flies were trained to enter and reenter a hole as the operant response. Moving in and out of the hole was detected with two infrared emitter and detector pairs. On each side of the hole, seven lines of light-emitting diodes (LED) were arranged in alternations of green and yellow. LED's were turned on when a session started and were turned off when the fly entered the hole. The reinforcer was sucrose solution delivered at the bottom of the hole by the needle of a glass syringe.

In Step 1, there is an assumed representation of behavior (rb) which elicits entering the hole [($rb \rightarrow ucr/R$)]. That *representation of behavior* (rb) becomes salient by being paired with the sucrose reinforcement ucs/s^{R+}. This pairing, [$rb \rightarrow ucr/R$] – ucs/s^{R+} is an Sensory or Motor Order 2 action.

In Step 2, the salient *representation of behavior* (rb) which elicits ($\rightarrow$) the operant response (ucr/R) is paired with the environmental stimulus (s). Here the operant behavior (ucr/R) is entering the hole which gets to the reinforcement (ucs/s^{R+}). This pairing of salient *representation of behavior rb* and environmental stimulus s, represented as s – [$rb \rightarrow ucr/R$], is an Sensory or Motor Order 2 action.

In Step 3, the environmental stimulus (s) is paired with the sucrose reinforcement (ucs/s^{R+}) making the environmental stimulus (s) more salient and valuable. This pairing acts to produce an incentive (Killeen, 1982a, 1982b, 1984; 1985). The environmental stimulus (s) takes on the elicitive properties of sucrose reinforcement ucs/s^{R+}. This is represented as s – ucs/s^{R+}.

Each of these steps on its own is a Sensory or Motor Order 2 action. The coordination of the three steps, on the other hand, is a Circular Sensory-Motor Order 3 task action.

Example 2. In this example, the three steps are illustrated using Schiller's (1949) study on *Octopus vulgaris*.

In a second example, *Octopus vulgaris*, the three steps of respondent conditioning are illustrated when Octopus vulgaris operantly conditions during maze learning. Two inverted cans, one covering a baited, the other an unbaited container was used. A partition wall had to be circumvented to reach the baited can. *Octopus vulgaris* learned to make a turn toward the proper side if the bait was visible all the time.

In Step 1, there is an assumed *representation of behavior* (rb) that elicits taking the detour by circumventing the partition wall (ucr/R). That *representation of behavior* (rb) becomes salient by being paired with the crab bait (ucs/s^{R+}). This pairing, [$rb \rightarrow ucr/R$] – ucs/s^{R+}, is a Sensory or Motor Order 2 action.

In Step 2, Schiller (1949) indirectly shows that the now salient *representation of behavior* (rb) which elicit the operant behavior (ucr/R) is paired with prior environmental stimulus (s), the visible bait can. Here operant behavior R is turning to the proper side to avoid the opaque wall and get to the baited can. The pairing of salient *representation of behavior* (rb) and environmental stimulus (s) is an Order 2 action. This is represented as s – [$rb \rightarrow ucr/R$].

In Step 3, the environmental s, the visible bait can, is paired with the crab bait (ucs/s^{R+}). This makes the s more salient and valuable. This pairing acts to produce an incentive (Killeen, 1982a, 1982b, 1984; 1985). The environmental s takes on the elective properties of ucs/s^{R+}. This is represented as s – ucs/s^{R+}.

Again, each of these steps on its own is an Sensory or Motor Order 2 action. Coordination of the three steps, on the other hand, is a Circular Sensory-Motor Order 3 task action.

Example 3. In this example, the three steps are illustrated using Andrew and Savage's (2000) study on *Lymnaea* (Pond Snail).

In a third example, *Lymnaea*, the three steps of respondent conditioning are illustrated when Octopus vulgaris operantly conditions during appetitive learning. *Lymnaea* was placed in a glass gutter. The gutter was placed within a white surround, 30 cm high. Halfway along the gutter, and visible through its sides, two panels, either black or white, were placed on either side of the gutter. *Lymnaea* were reinforced with sucrose when its head reached the level of the panels. *Lymnaea* learned to reach the level of panels, either black or white.

In Step 1, there is an assumed *representation of behavior* (rb) that elicits moving towards the level of the black and white panels (UCR/R). That *representation of behavior* (rb) becomes salient by being paired with the sucrose (UCS/s^{R+}). This pairing, $[rb \rightarrow UCR/R] - UCS/s^{R+}$, is an Sensory or Motor Order 2 action.

In Step 2, Andrew and Savage (2000) indirectly show that the now salient *representation of behavior* (rb) which elicits the operant behavior (R) is paired with prior environmental stimulus (s), the visible black and white panel. Here operant behavior (R) is moving towards the level of the black and white panels to get the sucrose. The pairing of salient *representation of behavior* (rb) and environmental stimulus (s) is an Sensory or Motor Order 2 action. This is represented as $s - [rb \rightarrow UCR/R]$.

In Step 3, the environmental (s), the visible black and white panel, is paired with the sucrose (UCS/s^{R+}). This makes the s more salient and valuable. This pairing acts to produce an incentive (Killeen, 1982a, 1982b, 1984; 1985). The environmental s takes on the elective properties of UCS/s^{R+}. This is represented as $s - UCS/s^{R+}$.

Each of these steps on its own is a Sensory or Motor Order 2 action. Coordination of the three steps, on the other hand, is a Circular Sensory Motor Order 3 task action.

Relationship among order 1, order 2 and order 3

This differentiation between these three types of learning is actually an old one (see Rescorla, 1988). In that paper, Rescorla states that the three most studied forms of learning, are: *a*) learning that involves exposure to a single stimulus (New Automatic Order 1); *b*) learning that relies on the relation between two stimuli (New Sensory or Motor Order 2) ; and *c*) learning that examines the relation between an organism generated response (R) and a stimulus s (Step 2 of Circular Sensory Motor Order 3). All we are showing is that these differ in their hierarchical complexity.

The difference between Order 1 action and Order 2 action is that, for Order 1 action, the endogenously salient unconditioned stimulus automatically elicits the unconditioned response. Organisms behaving at Automaticity Order 1 would be insensitive to outcomes except in an evolutionary sense. That is, consequences may be selected for in an evolutionary sense if the single action leads to survival and reproduction. Very primitive animals, such as single cell organisms, differentially respond to stimuli, for example, rejecting non-food items. However, such simple animals do not change their behavior because of its being paired with other stimuli or immediate environmental consequences, other than in terms of processes like habituation or sensitization.

» CONCLUSION

This is the first revision of the order and the corresponding stage sequence of the Model of Hierarchical Complexity since it was created in 1982. Although it may be difficult to remember the new numbers, the names have all stayed the same. What may be of interest is that the axioms and new information made it possible to do this revision. ■

REFERENCES

Abramson, C. I., (1994). *A primer of invertebrate learning: The behavioral perspective.* Washington, DC, US: American Psychological Association

Andrew, R. J. & Savage, H. (2000). Appetitive Learning Using Visual Conditioned Stimuli in the Pond Snail, Lymnaea. *Neurobiology of Learning and Memory*, *73*(3), 258–273. doi:10.1006/nlme.1999.3933

Butterfield, N.J. (2001). Ecology and evolution of Cambrian plankton. In R. Riding (Ed), *The Ecology of the Cambrian Radiation* (pp. 200–216). New York: Columbia University Press.

Commons M. L., Miller, P. M., Commons-Miller, L., & Chen, S. (2012). *A comparative account of numerosity in human and non-human animals*. Unpublished manuscript.

Commons, M.L., & White, M.S. (2006/2009). *U.S. Patent No. 7152051*. Alexandria, VA: U.S. Patent and Trademark Office.

Eisenstein, E. M., Brunder, D. G., & Blair, H. J. (1982). Habituation and sensitization in an aneural cell: Some comparative and theoretical considerations. *Neuroscience and Biobehavioral Reviews*, *6*(2), 183-194. doi:10.1016/0149-7634(82)90054-9

Eisenstein, E. M., Eisenstein, D. D., & Smith, J. C. (2001). The evolutionary significance of habituation and sensitization across phylogeny: A behavioral homeostasis model. *Integrative Physiological & Behavioral Science, 36*(4), 251-265. doi:10.1007/BF02688794

Hamilton, T. C., Thompson, J. M., & Eisenstein, E. M. (1974). Quantitative analysis of ciliary and contractile responses during habituation training in Spirostomum ambiguum. Behavioral Biology, *12*(3), 393-407. doi:10.1016/S0091-6773(74)91601-0

Henderson, T., & Strong, P. N. (1972). Classical conditioning in the leech Macrobdella ditetra as a function of CS and UCS intensity. *Conditional Reflex*, *7*(4), 210-215. doi: 10.1007/BF03000220

Killeen, P. R. (1981). Incentive theory. *Nebraska Symposium on Motivation, 29,* 169-216.

Killeen, P. R. (1982). Incentive theory: II. Models for choice. *Journal of the Experimental Analysis of Behavior, 38*(2), 217-232. doi:10.1901/jeab.1982.38-217

Killeen, P. R. (1984). Incentive theory: III. Adaptive clocks. *Annals of the New York Academy of Sciences, 423*, 515-527. doi:10.1111/j.1749-6632.1984.tb23456.x

Killeen, P. R. (1985). Incentive theory IV: Magnitude of reward. *Journal of the Experimental Analysis of Behavior, 43*(4), 407-417. doi:10.1901/jeab.1985.43-407

Lawrence, M. A., Klein, R. M., & LoLordo, V. M. (2009). *Examining exogenous and endogenous mechanisms of temporal attention*. Presented at the 50th annual meeting of the psychonomic society, Boston, MA.

Mingee, C. M. (2013). Retention of a Brightness Discrimination Task in Paramecia, (P. caudatum). *International Journal of Comparative Psychology*, *26*(3), 202-212.

Mingee, C. M., & Armus, H. L. (2009). Unsuccessful reinforcement of a discrete action in paramecia, P. caudatum. *Psychological Reports*, 105(2), 533-538. doi:10.2466/pr0.105.2.533-538

Mpitsos, G. J., & Davis, W. J. (1973). Learning: Classical and Avoidance Conditioning in the Mollusk Pleurobranchaea. *Science*, *180*(4083), 317-320. doi: 10.1126/science.180.4083.317

Palkovits, M., & Záborszky, L. (1977). Neuroanatomy of central cardiovascular control. nucleus tractus solitarii: Afferent and efferent neuronal connections in relation to the baroreceptor reflex arc. *Progress in Brain Research*, *47*, 9-34. doi: 10.1016/S0079-6123(08)62709-0

Patterson D. J., (1973). Habituation in a protozoan Vorticella convallaria. *Behaviour*, *45*(3), 304-11

Pavlov, I. P. (1927). *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex*. Oxford, England: Oxford Univ. Press.

Purves, D. (2004). *Neuroscience*. Third edition, Massachusetts: Sinauer Associates.

Rescorla, R. A. (1988). Behavioral studies of Pavlovian conditioning. *Annual Review of Neuroscience*, *11,* 329-352. doi: 10.1146/annurev. ne.11.030188.001553

Schiller, P.H. (1949). Delayed detour response in the octopus. *Journal of Comparative and Physiological Psychology*. *42*(3), 220-225. doi:10.1037/h0056879

Sokolowski, M. B., Disma, G., & Abramson, C.I. (2010). A paradigm for operant conditioning in blow flies (Phormia terrae novae Robineau-Desvoidy, 1830). *Journal of the Experimental Analysis of Behavior*, *93*(1), 81-89.

Thompson, R., & McConnell, J. (1955). Classical conditioning in the planarian, Dugesia dorotocephala. *Journal of comparative and physiological psychology, 48*(1), 65. doi:10.1037/h0041147

BEHAVIORAL DEVELOPMENT BULLETIN

The model of hierarchical complexity as a measurement system

Michael Lamport Commons[1], Robin Gane-McCalla[2], Cory David Barker[3], and Eva Yujia Li[4]

1 Harvard Medical School
2 Dare Institute
3 Antioch University
4 Harvard Graduate School of Education

ABSTRACT

The model of hierarchical complexity (MHC) is a mathematical model based on the "Theory of Measurement" that has gone through a number of iterations as a measurement system (Commons, Goodheart, Pekker, et al., 2005; Commons & Pekker, 2008; Commons & Richards, 1984*a*, 1984*b*; Commons, Trudeau, Stein, et all, 1998). It sets forth the measurement system by which actions are put into a hierarchical order and each order is assigned an ordinal number. In this paper, the components of the model will be described: actions and tasks, measurement and operations, and the axioms, followed by an articulation of emerging properties from axioms, and then a description of orders of hierarchical complexity of tasks. These are a reworked smaller set of axioms, which are more measurement-theoretical in nature. They also parallel the informal conditions underlying the kind of complexity that the MHC entails.

KEYWORDS: model of hierarchical complexity, theory of measurement, stages, actions, tasks, subtasks, subsubtasks, nominal scale, operations, measurements

The model of hierarchical complexity (MHC) is a mathematical model that sets forth the measurement system by which actions are put into a hierarchical order. The model is used as a general, unidimensional developmental measure of difficulty across domains. Dawson-Tunik's (2006) studies have found that the stage of development scored according to the model of hierarchical complexity was consistent with multiple other instruments that were designated to score development in specific domains.

The model of hierarchical complexity is not the only theory of development based on task complexity. Other metrics of task complexity have been proposed as well. horizontal or classical information complexity is one of them. It describes the number of "yes-no" questions. In classical information complexity, if a task requires one such question, the answer would transmit 1 bit of "horizontal" information. Similarly, if a task requires two such questions, the answers would transmit 2 bits. Each additional 1-bit question would add another bit. horizontal complexity, then, is the sum of bits required by tasks that require "yes-no" questions. The number of actions is 2^n.

Older metrics of task complexity such as the horizontal complexity and others have a few limitations. What is promising about the model of hierarchical complexity is that it is a newer model which overcomes those limitations as it is not content bound, does not miss stages and does not have any assumptions.

It is based on vertical complexity and involves hierarchical information. Hierarchical complexity refers to tasks that require the performance of lower-level subtasks in order to perform more complex, higher level tasks.

The most important advantage of the model of hierarchical complexity is that there is only one sequence of order of hierarchical complexity of tasks in all domains (Theorem 4, Commons, Trudeau, et al., 1998). The model is applicable to any domain of development in both humans and animals, such as social, cognitive, personal and such. MHC also seems to have advantage over previous proposals about developmental stages of humans. While previous models attribute behavioral changes across a person's age to the development of mental structures, MHC posits that task sequences of task behaviors form hierarchies that become increasingly complex. According to this model, less complex tasks must be completed and practiced before more complex tasks can be acquired. Thus, it accounts for developmental changes. Furthermore, previous theories of stage have confounded the stimulus and response in assessing stage by simply scoring responses and ignoring the task or stimulus. The model of hierarchical complexity separates the task or stimulus from the performance. The participant's performance on a task of a given complexity represents the stage of developmental complexity. Another factor which sets this model apart from previous models is that it not only extends developmental stages up to 17 stages, but also includes subtasks and subsubtasks which explain what happens between those stages.

As explained above, the model has many advantages. It has gone through a number of iterations as a measurement system (Commons, Goodheart, Pekker, et al., 2005; Commons & Pekker, 2008; Commons & Richards, 1984a, 1984b; Commons, Trudeau, Stein, et al., 1998). The model's empirical usefulness has also been set forth in earlier papers (e.g. Commons, Goodheart, Pekker, et al., 2005; Commons, Rodriguez, Adams, Goodheart, Gutheil & Cyr, 2006; Commons & Pekker, 2008). However, the newest version of this model has never been formally described. In this paper, the components of the model will be formally described: actions and tasks, measurement and operations, and the axioms. The previous eight axioms have been revised and collapsed into four axioms. A new, fifth axiom has also been added. The paper also articulates the emerging properties from axioms and a description of stages. Additionally, the paper includes a description of *Theorem 4,*which shows that there is only one sequence of order of hierarchical complexity of tasks in all domains.

Distributivity as an example

Hierarchical complexity can be illustrated with the example of the distributive property. The distributive property refers to when two sides of an equivalence "=" are represented differently, yet are equal. The distributive property describes a characteristic feature of some binary operators, namely that one argument must be "distributed" to the various elements of the other argument. Take for example $a \times (b + c) = (a \times b) + (a \times c)$. That says that one distributes the $\times$ across each term connected by the $+$ action. The distributive property plays a fundamental role in more general contexts, such as the complex numbers and the definition of rings in modern algebra.

The distributive law serves as a motivation for a newer form of complexity, called hierarchical complexity, formally presented here. In the case of evaluating $a \times (b + c)$, the task of distributing is more hierarchically complex than the two-part task of first evaluating $b + c = d$ and then evaluating $c \times d$. In the case of $(a + b) + c$, the organization of two actions of addition is arbitrary and no more hierarchically complex than addition in the evaluation of $(a + b) + c$ or $a + (b + c)$, because addition is associative. In the case of evaluating $a \times (b + c)$ it is more hierarchically complex than the task of evaluating $(a + b) + c$, because evaluating $a \times (b + c)$ requires the two actions of addition and multiplication to be performed in a certain order.

Actions and tasks

In the context of the model of hierarchical complexity, *actions* are defined as behavioral events that produce outcomes. Actions may be attributed to organisms, social groups, and computers. Actions may be combined to produce new, more complex actions (Binder, 2000). *Events* are perturbations that can be detected by at least two independent paths (Commons, 2001). A *task* can be defined as a set of required actions that obtain an objective, though the performed actions may or may not complete a given task. The study of tasks appears in psychophysics, a branch of stimulus control theory in psychology (Green & Swets, 1966; Luce, 1959) and in artificial intelligence (Goel & Chandrasekaran, 1992).

Hierarchical structure of tasks

The hierarchical order is constructed by seeing how one action is more complex than another, as illustrated above with distributivity as an example. A higher order action is defined in terms of two or more order actions of one order below, and the higher order action non-arbitrarily organizes those next lower order actions, as illustrated by Figure1 below. Mathematically speaking, we refer to distribution used in long multiplication, such as $a \times (b + c)$, as organizing the lower order actions of addition and multiplication, in non-arbitrary ways.

By definition, only the coordination of two or more next lower order actions produces an action at the next higher order. Coordinating actions of different orders result in other types of actions, and observations of these differences allow descriptions to be given about the orders. Orders have subtasks and subsubtasks between them. *Subtask actions* organize only one action from the same order and one or more from previous orders. They are prerequisites to other same order tasks. For example, the seventh order of hierarchical complexity is called primary. At this stage, the ability to do simple logical deduction, and simple arithmetic is attained. Examples of tasks it accomplishes are counting, addition and multiplication. Counting is one subtask action that is a prerequisite for addition. Addition, is another subtask action, and is a prerequisite for multiplication. They do not coordinate two or more actions, but coordinate one action from the same order and one or more from lower orders. Such coordination does not result in an increment of order. *Subsubtask actions* coordinate actions from different orders that are precursors but not prerequisites for subactions (see Figure 1).

Orders of hierarchical complexity form an ordinal scale with the first four axioms and definitions that follow. A fifth axiom makes all of the orders of hierarchical complexity equally spaced—that is, of equal difficulty.

In the next section of the paper, a description of the mathematical basis that defines the model of hierarchical complexity will be given, followed by the presentation of the formal, axiomatic version of the theory.

Measurement and operation

Measurement is the process of associating numbers with entities or objects. In this section of the paper, a description of the components of the model will be given—the system of entities, concatenation and comparison mathematical operators and the assignment function.

System of entities. To develop a system of measurement, one must represent the entities to be measured. In this case, the entities are task actions of organisms, social groups, and computers (Krantz, Luce, Suppes, & Tversky, 1971). In the modern algebraic representation of the model of hierarchical complexity, actions are represented by letters or numbers. A system of entities, as a set of actions, is represented by letters such as *A*. Unless these actions are the most simple and irreducible of actions, they are composed of other actions concatenated together—the simplest actions do not act upon other actions.

Concatenation operators. In mathematics and logic, an operation is an agent which executes on one or more input values and produces a new value. Concatenation operators are methods in which actions are connected. They are represented by "∘". They specify the order in which actions are executed, so that the order is fixed and not commutative: $a \circ b \neq b \circ a$. In the simplest terms, the concatenation operator is an ordering relation on the entities and also may be described by order relations: $A = (a, b) = \{a, \{b\}\}$, an ordered pair.

It was stated that mathematically speaking, we refer to distribution used in long multiplication as organizing the lower order actions of addition and multiplication in a non-arbitrary way. The non-arbitrary organization of addition and multiplication is accomplished by the concatenation.

Comparison operators. It was stated that actions are put into a hierarchy. Higher order actions are defined in terms of next lower order actions and non-arbitrarily organize the next lower order actions. The comparison operator, represented by >, is used to arrange actions in a hierarchy.

In the case of the real numbers, the system of entities is ℝ, the real numbers, the comparison operator is > and the concatenation operator is +. In the case here, the entities are the actions in a system and the numerical relational structure is the ordinal numbers (positive numbers and zero). The comparison and concatenation operators are the same as they are for the real numbers.

Assignment function. The assignment function is used to assign a numerical relational structure to the complexity of actions, which allows the complexity of actions to be stratified hierarchically into orders. In other words, the assignment function numbers the incrementally increasing complexity of actions *as orders*. It assigns numbers to those actions based on the complexity of those actions. Mathematically, the assignment function is represented by the Greek letter φ (phi).

An assignment function is a *homomorphic mapping* that transforms the entities to be measured into a numerical relation structure. Abstract algebra studies sets that are endowed with operations that generate interesting structure or properties on the set. Therefore, functions that preserve the operations are especially important. These functions are known as *homomorphisms*. In our case, the function maps a set of actions and their concatenation to a number, n (positive whole number or zero). The numerical relational structure preserves the relationship between actions – the more hierarchically complex the actions, the higher the number assigned.

For example, if a is an n order action the assignment function φ assigns the number n to a, which is denoted by $\varphi(a) = n$. The assignment function, $\varphi(a)$, denotes the *order of hierarchical complexity* (OHC) of a task action. The *measure of hierarchical complexity at order n* is defined as the minimum number of simple actions required to complete an action of that order.

The most irreducible action is at order 0. Order 0 is not really an action in the usual sense because it is not planned or controlled by the machine but by the programmer. It has no order of hierarchical complexity and therefore cannot be reduced. The first order that has an order of hierarchical complexity is of order 1 actions. Order 1 actions do not organize any actions so they are simple actions, because those actions have no order; that is why

it is called order 0. The actions of order 2 are made out of actions of order 1. The actions of order 3 are made out of actions of the order 2, and so on. The repeating process of an order of actions defined in terms of next lower order actions produces the numerical relation structure, and stratifies orders of hierarchical complexity.

» AXIOMS

The measurement system of the model of hierarchical complexity is composed of axioms. Axioms are rules that are followed to determine how the model of hierarchical complexity orders actions to form a hierarchy. There are five axioms: *well ordered, transitive, chain rule, coordination rule* and *equal spacing* (optional axiom). The axioms presented in the sections that follow are simplifications, reductions, refinements, and improvements that build on Piaget (e.g., Inhelder & Piaget, 1958) and his intellectual descendants (e.g., Campbell, 1991; Campbell & Bickhard, 1986; Commons & Richards, 1984a, 1984b; Commons, Richards & Kuhn, 1982; Tomasello & Farrar, 1986).

The *well ordered axiom* and the *transitive axiom* are rules that describe how the orders of hierarchical complexity are arranged. These axioms are found in most systems of measurement (Krantz, Luce, Suppes, & Teversky, 1971) including an ordinal a system of measurement.

Axiom 1 (Well-ordered). *If $a > b$ then $\varphi(a) > \varphi(b)$.*

Axiom 1 states that if one action is less complex than another action, then the assignment function, which gives a numerical order of hierarchical complexity to an action, must preserve the action's order in the hierarchy. In other words, under the conversion of actions into numbers by applying the mathematical assignment function φ, action a remains more hierarchically complex than action b. Breaking this rule would mean that the order of hierarchical complexity would not be uniform for all actions in which simpler actions are non-arbitrarily put into order by the model of hierarchical complexity.

Axiom 2 (Transitivity). *If $a > b$ and $b > c$ then $a > c$.*

Axiom 2 means that if action a is more complex than action b, and action b is more complex than action c, then action a is more complex than action c. In other words, the transitive property places the actions in a sequential hierarchical order. Breaking this rule would be similar to breaking Axiom 1, in that the numerical relational structure could not preserve the non-arbitrary sequence of orders of hierarchical complexity.

The following axioms regard how the concatenation operator affects the assignment function: the *chain axiom* and the *coordination axiom* describe the arbitrary and non-arbitrary character of the order of actions.

Axiom 3 (Chain rule). $\varphi(a \circ b) = \max(\varphi(a), \varphi(b))$ if $\varphi(a \circ b) = \varphi(b \circ a)$.

Axiom 3 states that when actions a and b are chained together in some order, and the order in which they are executed is not influential to accomplishing a task, the order of hierarchical complexity of $(a \circ b)$ equals to that of the highest subaction. Chaining together the two actions does not produce an action that is hierarchically more complex than either of the subactions.

Consider a scenario when a person's goal is to put on a pair of socks. It does not matter which sock a person puts on first, because the end result of the task is the same – both feet have socks on them. Another example would be a child who wants to try all the equipment at a playground. If the child sets out to play on swings and use the slide, it does not matter in which order these actions take place, so long as both are accomplished. The sequence in which the actions are combined does not bring about a higher order of complexity. In this case, ○ is chaining together of actions in some order. But with chaining, the hierarchical complexity of the new task does not increase. What this rule means, is that the way that two actions are combined is arbitrary. This is shown by the commutivity of a and b.

By axiom 3, an action of order n organizes at least two actions of order $n-1$, each of which in turn organizes at least two actions of order $n-2$, and so forth, until we reach the lowest-order, simple actions.

Axiom 4 (Coordination rule). $\varphi(a \circ b) = \max(\varphi(a), \varphi(b)) + 1$ if $\varphi(b) = \varphi(a)$ and $\varphi(a \circ b) \neq \varphi(b \circ a)$.

In this case, ○ coordinates the organization of the ordering of action rules in a non-arbitrary way. In addition, action a and action b has to be on the same stage. When these two conditions are satisfied, the coordination of action a and action b, which is represented by $(a \circ b)$, is one order more complex than either of the subactions. $\varphi(b) = \varphi(a)$ is necessary because, in order for the coordinated action to move up a stage, the actions have to be on the same stage. The coordination of two actions on different stages does not produce an action that is one stage higher.

A permutation, P, can be defined as the same elements happening in different orders, for example (x, y) or (y, x). Such permutations are not commutative of this axiom, i.e., $(x, y) \neq (y, x)$. With the *coordination axiom*, not all permutations, P, are acceptable, that is, not $P(a, b)$. This rule indicates that there is a possible match between the model-designated orders and the real world functioning of the order of those actions which the model-designed orders represent. To give examples, consider the above two. Returning to the sock example, one does not put shoes on first, then put socks over the shoes. Similarly with the child at the playground, the child must climb the stairs to the slide before going down the slide. These are examples of coordinating actions.

Theorem 4. There is only one sequence of order of hierarchical complexity of tasks in all domains (Commons, et al., 1998.) Axioms 3 and 4 together form the basis for Theorem 4.

Domain and order. Historically, stage theories utilized the notion that tasks at different orders within the same domain are "qualitatively different" (Kohlberg & Armon, 1984). Order of a task within a domain is determined by applying Axioms 3 and 4. To determine the order of hierarchical complexity of a task, one must count the number of actions in the hierarchy leading up to the task-required action. Each action in the hierarchy organizes actions from the previous order. Axioms 3 and 4 describe the conditions under which an action will be vertically higher in order than actions from the previous order. By checking repeatedly to see if an action requires a previous order action which in turn requires a previous order action, one can determine the linear hierarchy for a single sequence of tasks.

The order numbers describe the same order of hierarchical complexity of task-required actions irrespective of domain. Thus, one can map any developmental order sequence into the model of hierarchical complexity. This result may not directly imply synchronous development but suggests its possibility (See Axiom 5 below). Whereas the order numbers may be the same, the orders of performance may develop at different times. From an analytic perspective, the task requirements are constant and unvarying for different individuals regardless of how the subject performs on the task. The order complexity of each task within a sequence of tasks can be directly compared to the order of complexity for another set of tasks. The non-order of complexity aspects of tasks only makes it more difficult to apply. Stage systems must propose a parallel set of stages or levels. This parallelism is exemplified by the work of Fischer (1980), and Campbell and Bickhard (1986), and by the work of others including Armon's (1984) ethical development, Colby and Kohlberg's (1987) moral development, and Loevinger's ego development (Loevinger & Blasi, 1976; Loevinger & Wessler, 1970) as systematized by Cook-Greuter (1990) and Pascual-Leone's (1984) organismic theory of life orders.

Axiom 5 (Equal spacing[1]). $OHC(n + 1) - OHC(n) = 1$.

Here, $OHC(n) = \varphi(a)$, then for every order n, $(n)(OHC(n + 1) - OHC(n)) = 1$.

Axiom 5 states that the a priori difficulty of a task action changes by 1 for each change in the order of hierarchical complexity, irrespective of what adjacent orders of hierarchical complexities one is comparing. In other words, there is equal spacing between each order. This implies that the difficulty of going to the next order is the same regardless of what order is being required. This allows one to treat orders as actual numbers, and not just indication of relative position.

It might mean the order of hierarchical complexity, n, is a measure of the quantity of hierarchical information. Given that tasks at order $n + 1$ are defined by and coordinate two or more tasks at order n, the minimum number of order 1 tasks that an order n task is 2^n. Equal spacing might indicate that 2^n is well defined and therefore, $\log 2^n = n$, a parallel notion to bits. That might mean that n is a measure of the quantity of hierarchical information and could be called $Hbits$.

» A FORMAL DEFINITION OF THE MODEL OF HIERARCHICAL COMPLEXITY

With the above mentioned, we can now give more specific definitions about the model of hierarchical complexity. There are certain properties that emerge when certain rules are in play. In this section, these properties are articulated as definitions.

Definition 1. *There exists simple actions, x with $\varphi(x) = 1$. This is the lowest order action.*

Definition 2. *If there is no action, such as a computer calculating what has been programmed by a person, then the null action, such as the computer action, is at order 0. The computer program may act*

1. Optional

at a higher order, but it is just a reflection of a programmer solving a very much higher order task. There is no flexible action. That does not mean that variables may not be part of the program, or that randomness could not also be generated, but that the program only does what it is programmed to do. When a person enters information to a computer and a program, the action is by the person. All the action is done under control of the programmer. The exception would neural networks and stacked neural networks and the like.

Definition 3. *A higher order hierarchically complex action is defined in terms of two or more next lower order actions, $A = \varphi(B \circ C)$, where $\varphi(B)$ and $\varphi(C)$ are both less hierarchically complex than $\varphi(A)$ if $\varphi(B \circ C)$ is a coordination. This creates the hierarchy:*

> *Given $A = \{a, b\}$, we have that a and b are lower order of hierarchically complexity than A and together compose the set A.*

> *$A \neq \{A, …\}$ since a set cannot contain itself (See Russell's paradox (1902; 1980).*

This definition follows definition 1 and Axiom 4.

Next, the differences between chain rules and coordination rules are explained in more depth.

Definition 4. *Given a permutation of concatenated actions $R = (i_1, i_2, …, i_n)$ of the natural numbers 1, 2, …, n, the execution of action A is simply $A_{i_1} \circ A_{i_2} \circ … \circ A_{i_n}$.*

The rule R is a chain rule if the outcome of the action is the same for all $n!$ permutations of the numbers 1, 2, …, n. The outcome of the order of actions, $A_{i_1} \circ A_{i_2} \circ … \circ A_{i_n}$ is the same for all permutations $(i_1, i_2, …, i_n)$ of 1, 2, …, n.

Rule R is a coordination rule if there exists at least one permutation of actions $R = (j_1, j_2, …, j_n)$ of the numbers 1, 2, …, n so that the execution of the actions A_i i.e., $A_{j_1} \circ A_{j_2} \circ … \circ A_{j_n}$, is *not* the same as the outcome of the action A. Hence, the outcome of A_i is given by at least one, but *not all*, permutations of the A_i. This extends similarly to the cases where A consists of infinitely many actions.

Note that by Axiom 4, a coordination action $A = (\{A_1, …\}, R)$ necessarily coordinates subactions of subtasks of equal orders of hierarchical complexity (i.e., $\varphi(A_1) = \varphi(A_2) = …$). Thus the order of hierarchical complexity of A is one higher than the order of hierarchical complexity of all its subactions. Therefore, A_1 may be replaced by any subaction of A and still obtain the same result. As a consequence of these axioms, we see that if we let $\mathcal{A}$ denote the collection of all actions in a given system, then the order of hierarchical complexity is a function $h: \mathcal{A} \rightarrow \mathbb{N}$, where $\mathbb{N} = \{0, 1, …\}$ is the set of natural numbers (and zero) under the usual ordering.

The following properties emerge from the axioms and the definitions:

1. *Discreteness*: The order of hierarchical complexity of any action is a nonnegative integer. In particular, there are gaps between orders.

2. *Existence*: If there exists an action of order n and an action of order $n + 2$, then there necessarily exists an action of order $n + 1$.

3. *Comparison*: For any two actions A and B, exactly one of the following holds: $\varphi(A) > \varphi(B)$, $\varphi(A) = \varphi(B)$, $\varphi(A) < \varphi(B)$. That is, the orders of hierarchical complexity of any two actions can be compared.
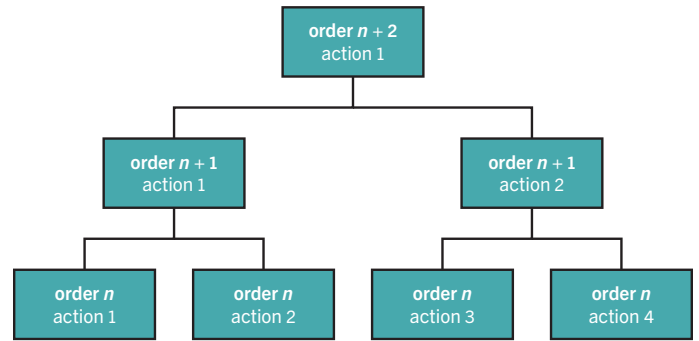
Figure 1. Hierarchical structures of tasks

4. *Non-reducibility*: A higher order action cannot be equal to any lower order actions. This property arises from the coordination rule, which claims that the coordination of two or more actions at the same order produces an action that is one order above.

Concepts from set theory are applied here to clarify why two order tasks can be non-arbitrarily ordered only at the next order. The higher order corresponds to a set A. Assume $A = \{a, b\}$. The lower order relations in the system correspond to the elements of lower order elements of the set, actions a and b. This creates the hierarchy because the set A is not the same as its elements a and b. The elements are at a lower order than the set. Therefore, the order of the set is not equal to the order of its elements, and $n + 1 \neq n$. Hence, the orders cannot be collapsed.

For example, consider an empty set $\emptyset$. Russell argued that an empty set cannot be a member of itself (Godehard, 2004). An empty set $\emptyset = \{\ \}$ has no member. Having no members means that there is nothing in it, or the member is "nothing". Because $\emptyset$ is a set, it is "something". Something cannot equal to nothing. Therefore, an empty set $\emptyset$ cannot equal to its member. Likewise, a higher order action cannot equal to any lower order action from which it is made.

This is consistent with Inhelder and Piaget and the model of hierarchical complexity. These theories state that each next order actions coordinates the actions performed at the preceding order of complexity. To apply the premise successfully, the actions of each stage must be unambiguously specified. The stage generator concept successfully eliminates ambiguity about what makes a stage a stage by precise specification.

Stages defined

The notion of stages is fundamental in the description of human, organismic, and machine evolution. Previously it has been defined in some ad hoc ways; here it is described formally in terms of the model of hierarchical complexity. Given a collection of actions $\mathcal{A}$ and a participant s performing $\mathcal{A}$, the *stage of performance* of S on $\mathcal{A}$ is the highest order of the actions in $\mathcal{A}$ completed successfully.

$$\text{Stage}(s, \mathcal{A}) = \max\{h(A) \mid A \in \mathcal{A} \text{ completed successfully by } s\}$$

This is in agreement with previous definitions (Commons, Trudeau et al. 1998; Commons and Miller 2001).

» CONCLUSION

This paper accomplishes two goals. First, it formalizes the Model of Hierarchical Complexities as a measurement system. The components of the measurement system are conceptualized for the first time—the system of entities, comparison operator, concatenation operator and the assignment function. The comparison and concatenation operators describe how the actions are structured. The assignment function is the procedure which assigns numbers to the actions.

Secondly, this paper clarifies key concepts of the model. Axiom 4 articulates the coordination rule. The 4th property, non-reducibility, is demonstrated by proving that a higher order action cannot equal to any lower order actions. In addition, the chain rule and coordination rule are put into mathematical expressions, making them succinct and absolute. This paper substantializes axioms and definitions of the model, which provides theoretical foundation for utilizing the model to measure the task order of actions. ∎

REFERENCES

Armon, C. (1984). Ideals of the good life and moral judgment: Ethical reasoning across the life span. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), *Beyond formal operations:Vol.1. Late adolescent and adult cognitive development* (pp. 357–380). New York: Praeger.

Binder, C. (2000). *Component/composite analysis and programming from the real world*. Paper presented at the annual meeting of the Association for Behavior Analysis, May, in Washington, DC.

Campbell, R. L. (1991). Does class inclusion have mathematical prerequisites? *Cognitive Development, 6*(2), 169-194. doi:10.1016/0885-2014(91)90035-C

Campbell, R. L. & Bickhard, M.H. (1986). *Knowing levels and developmental stages*. Basel, Switzerland: Karger.

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

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

Cook-Greuter, S. R. (1990). Maps for living: Ego-development theory from symbiosis to conscious universal embeddedness. In M. L. Commons, J. D. Sinnott, F. A. Richards, & C. Armon (Eds.), *Adult Development: Vol. 2. Comparisons and applications of adolescentand adult development models* (pp. 79–104). New York: Praeger.

Commons M.L. (2001). Problems with mentalistic explanations, freewill, self, soul, and intrinsic motivation. *Metanexus: The Online Forum on Religion and Science*, 08.21. http://www.metanexus.net/magazine/ArticleDetail/tabid/68/id/3966/Default.aspx

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

Commons M.L., Goodheart, E.A., Pekker, A., Dawson-Tunik, T.L., Cyr, E., Rodriguez, J.A., et al. (2005). The relationship between orders of hierarchical complexity and Rasch scaled stage scores: Balance beam, laundry,and counselor-patient task sequences. *Journal of Applied Measurement,10*(2), 1-11.

Commons, M., Rodriguez, J., Adams, K., Goodheart, E., Gutheil, T., & Cyr, E. (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*(6), 430-435.

Commons, M. L., Trudeau, E. J., Stein, S. A., Richards, F. A. & Krause, S. R. (1998). The existence of developmental stages as shown by the hierarchical complexity of tasks. *Developmental Review 8* (3), 237-278.

Commons, M. L., Richards, F. A. & Kuhn, D. (1982). Systematic and metasystematic reasoning: A case for a level of reasoning beyond Piaget's formal operations. *Child Development, 53*(4), 1058-1069. doi:10.2307/1129147

Commons, M. L. & Miller, P. M. (2001). A quantitative behavioral model of developmental stage based upon hierarchical complexity theory. *Behavior Analyst Today, 2*(3): 222-240.

Commons, M. L., & Pekker, A. (2008). Presenting the formal theory of hierarchical complexity. *World Futures: Journal of General Evolution* 65(1-3), 375-382. doi:10.1080/02604020802301204

Commons, M. L., & Richards, F. A. (1984*a*). A general model of stage theory. In M. L. Commons, F. A. Richards,& C. Armon (Eds.), *Beyond formal operations: Vol. 1. Late adolescent and adult cognitive development* (pp. 120-140). NY: Praeger.

Commons, M. L., & Richards, F. A. (1984*b*). Applying the general stage model. In M. L. CommonsF. A. Richards, & C. Armon (Eds.)*, Beyond formal operations: Vol. 1. Late adolescent and adult cognitive development* (pp. 141-157). NY: Praeger.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review, 87*(6), 477–531. doi:10.1037/0033-295X.87.6.477

Godehard, L. (2004). *One hundred years of Russell's paradox*. Germany: Walter De Gruyter Inc

Goel. A. & Chandrasekaran, B. (1992). Case-based design: A task analysis. In C. Tong and D. Sriram (Eds.), *Artificial intelligence approaches to engineering design, Vol. II: Innovative design*, (pp.165-184). San Diego: Academic Press.

Green, D. M. & J. A. Swets. (1966). *Signal detection theory and psychophysics*. Huntington, NY: Krieger.

Inhelder, B. & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic Books.

Kohlberg, L., & Armon, C. (1984). Three types of stage models used in the study of adult development. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), *Beyond formal operations: Vol. 1. Late adolescent and adult cognitive development* (pp. 357–381). New York: Praeger.

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

Loevinger, J. (1970). *A theory of ego development.* San Francisco: Jossey–Bass.

Loevinger, J., & Blasi, A. (1976). *Ego development: Conceptions and theories.* San Francisco: Jossey–Bass.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New York: Wiley.

Luce, R. D. (1963). Detection and recognition. In R. Luce and R. Bush (Eds.), *Handbook of mathematical psychology*, (pp.103-90). New York: Wiley.

Pascual-Leone, J. (1984). Attention, dialectic, and mental effort: Towards an organismic theory of life stages. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), *Beyond formal operations: Vol. 1. Late adolescent and adult cognitive development* (pp. 182–215). New York: Praeger.

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

Russell, B. (1902). Letter to Frege. In van Heijenoort, J (Eds.), *From Frege to G*ödel, (pp. 124-125), Cambridge, MA : Harvard University Press.

Russell, B. (1980*). Correspondence with Frege. In Philosophical and Mathematical Correspondence,* by GottlobFrege. Translated by Hans Kaal. Chicago: University of Chicago Press.

Tomasello, M. & Farrar, M. J. (1986). Joint attention and early language. *Child Development 57*(14), 54-63.

What are the relationships between four notions of stage change?

Michael Lamport Commons

Harvard Medical School

ABSTRACT

There have been a number of models for transition between stages, including Piaget's dialectical model, Dawson's use of Rasch score values, and the newest, the systematization resulting from micro-genetic research. This paper discusses four different accounts of stage transition, each delineating how to obtain data on stage transition for each method. It also discusses how the methods might be interrelated. First, the paper elaborates the original Piagetian model by systematizing the transition substeps using choice theory and signal detection. An examination of stage transition included scoring interviews or other participant responses for statements that reflect each of these steps. Secondly, the paper examines micro-developmental approaches. These approaches identify what may be potential subtask and subsubtask actions that may occur during transition to the next stage. Then, the paper describes and illustrates the use of Rasch analysis to quantify the extent to which a participant's performance on an instrument is transitional. This approach might numerically pinpoint where in the transition an individual is but it did not measure the difficulty of the specific task subtask actions (strategies). A method for combining stage scores, subtask action scores, and the sub-subtask action scores was introduced. Finally, the paper presents a methodology for creating tasks and methods of support that directly measure transition. The purpose of this approach was twofold. First was to empirically test for the transition subtask and subsubtask actions extracted originally from the interview process. Second was to figure out how high in transition an action would get with support.

KEYWORDS: stage transition, Rasch analysis, micro-genetic research, micro-developmental approaches, subtask, subsubtask, dialectical theory

There are at least four different versions of stage transition. This paper will discuss what they are and how some of them might be interrelated. Stage transition was early outlined by Piaget in this dialectical theory of stage change. Piaget suggested a dialectical theory of transitional steps. To describe transition, this model elaborates on and systematizes the dialectical strategies described in the Piagetian probabilistic transition model (Flavell, 1963; Flavell, 1971). This systematization of the substeps was based on choice theory and signal detection (Richards & Commons, 1990).

Transition concepts are represented by a number of different forms of measurement. There are at least four different methods: The conceptions roughly fall into three groups. The first is how steps and strategies are seen in interviews, narratives, writing sample etc. The second is how the steps and strategies are seen from observing how a person works on a task during transition. The tasks are often problems to be solved. The third, is to give instruments that through cuing or some other means, shows what step the person can do with some support.

The four accounts of stage change are:

1. Scoring interviews directly for stage and steps.

2. Micro development: designing, constructing tasks, subtask actions, subsubtask actions and prompts that measure transition directly.

3. Rasch scaling and measuring steps, subtask actions, and sub-subtask actions (strategies) during transition.

4. Transition steps and substeps.

» SCORING INTERVIEWS DIRECTLY FOR STAGE AND STEPS

In the scoring of interviews and narratives, the interviews and narratives describe task solutions, and the scorer attempts to interpret in the statements how the interviews and narratives reflect the steps in transition. Ross (2008b) used this method in her dissertation. The Commons, et al. (2005) scoring manual has lots of examples of this. Also see Miller and Lee (2000, June); Miller, Lee, and Commons (2000, June).

Correspondence regarding this article should be addressed to Dr. Michael Lamport Commons, Harvard Medical School, 234 Huron Avenue, Cambridge MA 02138. E-mail: commons@tiac.net

After scoring the performances, such as by using a program like Dawson's The Lectical Assessment System (Dawson & Heikkinen, 2009; Stein & Heikkinen, 2008), or scoring by hand, one can perform a Rasch Analysis. Rasch (1960/1980) analysis scales performance and items on the same log linear line. Transitional performance is shown by the mixtures of performances at different stages. The mixtures range from 0% at the higher stage to 100%. We call 95% at a stage consolidated performance and 0% up to 95% transitional. The advantages of the Rasch analysis are that:

1. It reduces measurement variance to a minimum, and

2. Thus yields *direct* comparability.

The Rasch scores can be translated into stage score. Intermediate stage score show transition. For each part of transition, there are characteristic statements as Dawson has pointed out.

» MICRO DEVELOPMENT: DESIGNING, CONSTRUCTING TASKS, SUBTASK ACTIONS, SUBSUBTASK ACTIONS AND PROMPTS THAT MEASURE TRANSITION DIRECTLY

Informally, macro development refers to changes in behavior and the control of behavior that take place across stages. Micro development refers to changes that take place within or between stages. Micro Development is based on finding what subtask actions, and subsubtask actions (strategies) are used during the transition to the next stage. It is not fractal and they are task sequence dependent. The purpose of the micro development approach is twofold. First is to see what evidence there is for systematic change outside of the interviews process. The second is to systematize the strategies of individuals that one observes in micro-developmental studies. One would see what strategies they used in a number of ways. One could observe how they attack certain stage change examples. Also, one could reinforce correct answers.

By seeing what steps they can do when cued and what strategies they use at each of these steps, one learns what difficulty individuals may overcome and how those individuals overcome them. Note that support in this context is not for a whole stage but in the context of a transitional step between a stage. What is critical is to figure out how high a step they get to, with support. In summary, one could model the step strategy and one could cue it.

Commons-Miller and Commons realized there were subtasks and subsubtasks between the orders of hierarchical complexity. This was clear from Boom et al (in press) work on overlapping waves (See Siegler, 1996) shown with Rasch on stage transition. As inspired by Boom, et al (in press), we differentiate three additional levels of task actions: next higher order tasks, subtask action, and subsubtask action. A *next higher order task* requires that the new task action (*a*) be defined in terms of two or more next lower order subtask actions; (*b*) organize them; (*c*) in a non-arbitrary way. A *subtask action* is defined in terms of (*a*) only one same order action; and (*b*) another next order or even lower order action. It is not a next order task action because there is only one next lower order task action that the task action operates on. One subtask action is a necessary prerequisite for the next subtask action. A *subsubtask* action is defined in terms of one or more actions two orders down. So they really operate mainly on just one next

lower order action. Subsubtask actions may be sufficient but not necessary for the next subsubtask action. The sequence of the actions acquired often depends on the sequence that the teacher provides for the student. There is just one thing that organizes subtasks actions. That is, one subtask action serves as a prerequisite for the next subtask action. The subsubtask actions have a weaker relationship and may be arbitrary organized or one may serve as precursor, and may be only sufficient but not necessary for the next subsuborder task action. The sequence in which these actions are acquired often depends on the sequences that the teacher provides for the students. And it may require more horizontal complexity than the preceding task, as is the case of adding more than two numbers together.

For example, a primary order task is to coordinate preoperational task actions. A preoperational task action coordinates sentential task actions. Coordinating the preoperational order a sentential order task within a primary task action would be subtask.

The metric one might use as output would include the stage scores, subtask action scores, and the subsubtask action scores. These would be $x.y.z$, a three digit decimal that we would consider ordinal. The x = the stage score, the y would be the subtask action score and the z = subsubtask action score. So the first order task, first subtask action, and the first subsubset action would be 1.1.1. If one wants to add stage, it would be (10) 1.1.1 for example.

The three layers of task action may be illustrated with counting and arithmetic tasks. A subtask action within the primary stage is to learn to count. The subsubtasks actions would be the sequence participants go through in moving from preoperational counting of objects in a line to counting objects in a random array. One set of actions come from the preoperational order 7. This requires numbers to be said or indicated for each object in a line array. The problem is that saying numbers from the sentential Order 6 continues after running out of objects. In primary order 8, one keeps track of what one has counted. This is a prerequisite for the second subtask, which is to learn how to add. Adding is a prerequisite for the third subtask, learning how to multiply. Adding is a subtask action for doing distribution (Long multiplying), but it is not its own next order action. Likewise, multiplying is also a subtask. There are just three things that organize sub tasks actions. The strongest is a prerequisite, weaker is precursor, and the other is that there is more horizontal complexity.

Preoperational order 7 actions organize sentential order 6 actions

Like all orders in the Model of Hierarchical Complexity (Commons et al., 1998; and for an earlier and similar model, Fischer, 1980), we characterize Preoperational Order 7 actions by how they are defined, how organisms do the actions, and the end result. Preoperational Order 7 actions are defined in terms of two or more Sentential Order 6 actions. They organize the Sentential Order 6 actions. The organization is non-arbitrary.

For example, organisms form lists of organized set of acts. They make simple deductions that connect simple sequences of actions. Humans telling stories are like sequences 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, combine simple propositions, and make simple deductions.

Arithmetic at the preoperational order

Counting preordered objects is preoperational behavior. Some organisms or preschoolers count. They apply sentential stage ordinal sequences to novel sets of objects placed in a line. One Sentential order 6 task is: *a*) Saying numbers in the order they were taught; *b*) another sentential order 6 task is pointing to or touching all of a set of objects one by one. The first sentential action is acting out an ordinal sequence by saying the numbers including above the first few. We do not know exactly what the primate representational sequence is, but we do know they have some sort of representation that we will call magnitude (Gallistel & Gelman, 2005). The second sentential stage action is going along the objects in the line one by one. When these two are combined, this is preoperational counting. Specifically, children point to each object in order. Then they apply a number from the number sequence. That number goes up one number in the sequence as they point to a new object. The last count may be called 5, five, cinco, etc. This also indicates the size of the set. This is elementary counting. Because organisms completing tasks at this order relate two sequences together, they may say the sequence of numbers or use number symbols in a sequence, and indicate which object is currently being counted when items are already arranged in a line. Other kinds of sequences may be interrelated as well.

Subtasks at the preoperational order

Before moving to the primary order, one of the subtasks actions is learning the "tens" labels. Since this must be learned before counting of larger numbers may take place, it is a prerequisite. One of these subtasks actions is learning the "tens" labels. Since this must be learned before counting of larger numbers may take place, it is a prerequisite. Applying the coordinations of number representation to "any number" of objects is required. At first, one counts items, but one does not stop after all the items have been "counted" A subtask at the preoperational order is keeping track of what has been counted. But learning to stop when one runs out of things to count awaits the primary order.

Arithmetic at the primary order

Counting. At the Primary Order, two or more actions from the Preoperational Order are coordinated. The first subsubtask action is to count disordered objects that are the same. The next subsubtask action is to count disordered objects that are not the same. The last number counted indicates the size of the set. For example, for five objects, the size of the set would be "5". There are three major subtasks required at Primary Order 8 Counting. The first subtask is true counting. The second subtask is addition and their inverses. The third subtask is multiplication and their inverses.

This first subtask of *true counting* is made possible by the suborder task: *a*) Having a way of marking that an object has already been counted by such action as moving it into a separate pile. Primary Order 8 task actions organize Preoperational Order 7 task actions. The first Primary Order 8 subtask actions may organize counts of organized objects from the Preoperational Order 7 tasks and apply them to very large numbers of randomly organized sets of objects. This is done by not only using the counting of objects from the preoperational order, but keeping track of what has been counted,

which is also from the preoperational order. Within the "counting" subtask action, there are a number of subsubtask actions. The first subsubtask action is to count disordered objects that are the same. The second is to count disordered objects that are not the same. The third subsubtask in true counting is very large numbers with randomly organized sets of objects. Hence, children count 100's of objects as opposed to 10 to 12. They learn the subtask actions of addition, subtraction and then multiplication (Van der Ven, Boom, Kroesbergen, & Leseman, 2011) and their inverses. This can connect ordinality to cardinality.

Addition. Note that addition is the second subtask and only operates on one action from the primary order, counting. This subtask use of accurate counts is addition/subtraction. This is true adding because they are using symbolic markers to insure that they have counted an item. This can as easily and accurately be done with numbers greater than 10. There are subsubtasks in addition. The first subsubtask action is to learn that the quantity remains constant when two sets are put together. The second subsubtask action is to combine sets. In the first subsubtask action, one counts the second set by continuing the count of the first. In the second subsubtask action, counts of sets can be combined by using the results of the count of the total from subsubtask action one and then continuing the count with the next number. The third subsubtask action is to know what the total of the first set is and what the total of the second set is and then to simply last subsubtask action is to learn the addition table. The third subsubtask action is to add those two numbers together. In literate cultures, the addition facts are learned. The last subsubtask action is to learn the addition table.

Multiplication. Adding is a prerequisite for the third subtask multiplication. There are some subsubtask actions for multiplication. The first subsubtasks for multiplication at the primary order is to see numbers in fixed arrays. An array problem is a problem where there are many rows of items. For example, one could see two rows of three red checkers. This makes it possible to see that groups of numbers have a meaning as opposed to just being "a bunch". The second subsubtask for multiplication at the primary order is to count by multiples of a number. One learns to count by 2's, 3's, 10's etc. The third subsubtask for multiplication at the primary order is to learn the multiplication is repeated addition. This means learning that $2 + 2 + 2$ is the same 3×2 and both answers are 6. This completes learning multiplication is a ways of understanding groups of numbers. The last and fourth subsubtask for multiplication is to memorize multiplication facts. The point to memorizing math facts is to establish automaticity. With automaticity, one recalls facts instantly in lieu of counting on fingers or diagraming the problem counting strategies.

» RASCH SCALING AND MEASURING STEPS, SUBTASK ACTIONS, AND SUBSUBTASK ACTIONS (STRATEGIES) DURING TRANSITION

Rasch analysis was has been used to confirm the order of the hierarchical complexity of stimulus items or tasks (Commons, Goodheart, Pekker, Dawson-Tunik, Cyr, & Rodriguez, 2005). This has been useful in assessing the nature of the items used to measure performance; the possible natural number order of hi-

erarchical complexity of each item, and the corresponding stage of performances on each item. (Mislevy & Wilson, 1996; Spada & McGraw, 1985; Wilson, 1989).

To measure how strategies are distributed during transition, one can take Boom's et al (in press) approach which is to specify the tasks, subtasks and sometimes the subsub tasks by using Rasch Analysis to produce the overlapping waves from the Siegler's (1996) Overlapping Waves Model. This will inform whether or not the subtasks and subsubstaks actions are in the right order and whether they are distinct. Siegler introduced the Overlapping Waves Model as a metaphor to illustrate a typical sequence of increasing and decreasing use of strategies during development. Those strategies may address tasks, subtasks, and subsubtasks. Boom, et al (in press) go beyond metaphor using Item Response Theory (IRT) (Ostini & Nering, 2006), to analyze such categorical longitudinal data.

Item response theory began with Frederic M. Lord (Lord & Novic, 1968), the Danish mathematician Georg Rasch (1960/1980), and Austrian sociologist Paul Lazarsfeld (1950; 1959). There are a number of people who furthered the progress of IRT (Andrich, 2004; Wright, 1990). IRT provides a framework for evaluating how well individual items on assessments work as well as overall assessments perform. Item response theory focuses on the theory on the item, as opposed to the test-level focus of classical test theory. Item might be multiple choice questions that have incorrect and correct answers or also statements on questionnaires that require participants to rate indicated level of agreement, or patient symptoms scored as present/absent. IRT is based on the idea that the probability of a correct/keyed response to an item is a mathematical function of person and item parameters. The person parameter is called latent trait or ability; it may, for example, represent a person's intelligence or the strength of an attitude. Item parameters include difficulty (location), discrimination (slope or correlation), and pseudoguessing (lower asymptote).

In Boom's analysis, strategy use is scored as an ordinal variable with few categories and longitudinal development as a vector of such scores. It provides the means to relate the use of such strategies to an underlying developmental dimension. Movement of individuals along this dimension can be modeled by means of Latent Growth Modeling. Latent growth modeling is a statistical technique used in the structural equation modeling (SEM) framework to estimate growth trajectory. It is a longitudinal analysis technique to estimate growth over a period of time. It is widely used in the field of behavioral science, education and social science. Latent Growth Models (Boom et al., 2001 Meredith & Tisak, 1990; Rao, 1958; Scher et al., 1960) represent repeated measures of dependent variables as a function of time and other measures.

» TRANSITION STEPS AND SUBSTEPS

Transition steps are somewhat different from transition subtask actions and especially subsubtask actions. For a review of the history, see Commons and Richards (2002) and Ross (2008a). They are more process oriented. They are fractal. Commons and Richards (2002) embellished on Piaget dialectical stage change notion combining Kuhn phases of transition with Piaget's dialectical steps that are also fractal (see Table 1).

So like a dynamical system, an increase in rate of reinforcement many be caused by small perturbations in reinforcement, such as at time b an additional reinforcer being earned. Once perturbed, switching to a new behavior may increases the rate further if it is further along in transition. But, switching is not deterministic, it is probabilistic.

What happens with the effect of reinforcement of switching more often, is that the rate of switching between A and B goes up. For example, the relativistic step, the action of switching is reinforced more often for doing A in certain and B in others, and vice versa. So A occurs in certain situation and gets reinforced,

Table 1. Combining Kuhn phases of transition with Piaget's dialectical steps

Step	Substep	Relation	Name	Piaget	Dialectical form
0		A = A' with B'	Thesis		Extinction Process
1		A fails	Antithesis	New Step	Negation or Complementation, Inversion or alternate thesis
2		B (or not A)	Relativism	Step 1	Alternation of thesis and antithesis depending on non-relevant context
3		A or B	Synthesis	Step 2	Random hits, false alarms and misses, correct rejections (Smash1)
4		A & B	Smash A & B together	New Step	Components from A and B are included in a nonsystematic, non-coordinated manner
	1		Random hits, false alarms, misses, correct rejections		Incorporates various subsets of all the possible components
5	2		Hits and excess false alarms		Incorporates subsets producing hits at stage n. Basis for exclusion not sharp (Overgeneralization)
6	3		Correct rejections and hits		Incorporates subjects that produce correct rejections at stage n but excess misses. Basis for inclusion not sharp (Undergeneralization)
7		A with B	Temporary equilibrium	Step 3	New temporary equilibrium

and likewise *B* can be also is reinforced in certain situations. The world is complicated and probabilistic enough that once in a while a behavior gets paid off. This is an example of melioration, in which the frequency of a response, *B*, relative to another response, *A*, increase as the rate of reinforcement for *B* increases.

The volatility of switching and forming combinations increases dramatically during smash. One cannot predict what combination will occur in first step of smash. One can says it is probabilistically what they are, but there are a variety of them. The interesting first step of smash before hits emerge is that there is so much variety. This is not a settled down deterministic system. Very small changes of reinforcement during smash, pulls it into the substep in which hits start to predominate. Finding the rate and acceleration of alternations of old-stage and newer-stage actions has never been tried. Finding the proportion of new-stage versus old-stage behavior has been found. A Rasch analysis is a more advanced form of this.

Speculated rate of change through the steps

There is no reinforcement gain in going to step 1, just an experience loss of reinforcement. Therefore, it might take a long time. There is very little gain at step 2 because the alternative does not usually work. There is slightly more at step 3. The flexibility does produce some gains in many cases. The big gains are at step 4, but it is also the most risky. Once into Step 4, progress should be fast because the acceleration of reinforcement as one gets hits is huge.

» HOW DOES STAGE CHANGE TAKE PLACE?

This is a brief summary of stage change interventions. Part of the issue is whether or not it is even possible to create generalizable stage change. There are some dilemmas or a paradox that Plato describes. Plato asserted that one cannot learn anything one does not already know. It is also Fodor's (1975; 2008) argument against learning. Becker (2001) has articulated that the Piagetian notion of reflection or reflective abstraction cannot be true. This is because it means one would already have to know the next stage behavior to reflect upon it. The solution has always been that the transition steps are always driven by gains in reinforcement. Sticking with the previous stage means a loss of potential reinforcement. Sticking which each step also means a loss of potential reinforcement.

» REFLECTIONS, METACOGNITION AND THE COMPRESSION OF STAGES DURING RECAPITULATION

Let us say that reflection is a form of metacognition. At first, reflection on one's performance costs an entire stage. So shadowing what one is doing is one stage higher than just doing the task. As lower stage actions become practiced, they become automatic. When they are automatic, they no longer take up as much "computational space" in M-space terms (Pascual-Leone, 1970; Pascual-Leone & Goodman, 1979). This would account for part of the compression we see in reflection. Even metacognitive actions can become automatic (Schrader, personal communication, 2005, June). The lower stages are integrated into the complexity of the higher stages, becoming one with the higher system as it were. For example, Commons does not think about variables at first when solving algebra problems. Usually, he only thinks of values of variables for

just long enough to move on to the variable along which they lie. So compression would be one form, and this rapid recapitulation of the stages would be a second form used in reflection.

But in Schrader (personal communication, 2010, June) and my theory of next stage performance, there is something like chunking (e.g. Gabriel & Mayzner, 1963; Gobet et al., 2001; Miyapuram, Bapi, Pammi, & Kenji, 2006) that transforms lower stage actions into new actions that do not require recapitulation of the lower stage actions. Even though the lower stages are integrated into the complexity of the higher stages, becoming one with the higher system as it were, they probably do not take up more "computational space." Unlike a computational model, however, the systems do not work like a model where the processes remain intact though are processed 'faster' or unconsciously in order to 'make room' for the higher more complex thought processes. They are chunked into units that take less space.

» STEP 0 AND STAGE CHANGE

The major problem with the reinforcement theory of stage change lies at step 0. What interventions might work and why? Without intervention and just exposure to the next stage task, the question is how does an organism know that there is more reinforcement ahead that they are not obtaining (Becker, 2001) without knowing about the next stage is some way? Bereiter (2006) considered 10 relatively neglected resources for the "bootstrapping" of cognitive growth, including chance. Here are some incomplete proposals that capitalize on chance.

1. The organism observes another organism obtaining more reinforcement on the same task. This is incomplete because there is always a first time of observing. Does the organism see the different response to the same task?

2. One possible mechanism is that the present stage action fails to obtain reinforcement period. On such an occasion, Step 0 could begin.

3. It might be such a local failure that it in itself in not overwhelmed by the relatively constant rate of reinforcement. Vaughan and Herrnstein (Herrnstein, & Vaughan, 1980; Vaughan, 1981; Vaughan, & Herrnstein, 1987) showed over and over that the most local rate of reinforcement is what controls behavior. Also, an organism might encounter a new problem for a lot of different reasons.

4. We do see stage change often when the environment abruptly changes. This is the case when gaining cross cultural experience (Commons, Galaz-Fontes, & Morse, 2006). The place to get stuck is before step 0.

There are three targets of stage change: individuals, groups including cultures, organizations and the like, and species. For human individuals, most of us hope that education will increase stage. Commons, Galaz-Fontes, & Morse (2006) found that the average moral stage was Concrete stage 9 in non-literate people in Baja California. Contrast this with mean stage of performance in the U.S. to be Formal stage 11 (Commons, 2008). Very preliminary data suggests that the more educational opportunity one provides, the greater the stage. But this is correlational data (Commons, Miller, & Kuhn, 1982). Already, the top 40% of the U.S. popula-

tion goes to college. What needs to be addressed is how to get individualized instruction and motivational techniques used with the rest of the population. But the distribution of stage has been highly resistant to modification above 1 stage increase which is a 2 stage in traditional moral development terms (Schlaefli, Rest, & Thoma, 1985). The degree of benefit for many interventions were studied by Grotzer, Commons, and Davidson (1986).

First, six forms of intervention for individuals are ranked by increasing levels of effectiveness as found by Grotzer, Commons, and Davidson (1986). This was on Inhelder and Piaget (1958) pendulum type problem.

1. Letting person create their own problems. This did the worst with people doing worse after intervention.

2. Practice without feedback on presented problems. There was a small but statistically significant effect.

3. Practice with feedback. This produced no improvement. This might be due to the fact that guessing resulted in such a great deal of punishment from finding out one was wrong half the time

4. 2 levels of support, given direct instruction, walking people through the task. Fischer (Fischer, Hand, & Russell, 1984; Fischer & Kenny, 1986) reports that this raises the stage of performance by 2.

   e) Whether or not this generalizes or is permanent is not known.

   f) 1 level of support, providing examples, providing a model. Fischer (Fischer, Hand, & Russell, 1984; Fischer & Kenny, 1986) has shown that this raises stage by one. With multiple examples, in related but different task sequences, this produces per meant and generalized performance change. This is the standard way to teach mathematics. The drawback is that people often quit the activity.

5. Reinforce correct answers. With reinforcement and feedback, they raised the stage of performance from concrete to formal in 75% of 5th and 6th graders. In a follow-up study, such performance did not generalize until the 7th grade.

» WHY SOCIAL AND ORGANIZATIONAL STAGE GOES UP

In the special issue of World Futures (2008), there are a number of papers that address this area (e.g. Commons, 2008; Commons, & Goodheart, 2008; Commons, & Ross, 2008; Glock-Grueniech, 2008; Inglis, 2008; Koplowitz, 2008; Robinett, 2008; Ross, 2008b; Ross, & Commons, 2008) Therefore it will not be discussed here.

» DISCUSSION

As Sara Ross might say, here we have four systems of viewing stage transition and no real unification. Yes, Rasch analysis may be applied to make sure the sequences of steps, scoring or subtasks and subsubtasks is correct. But that does not fit them into a single system. One way to view this, is that the systems address different aspects of transition and use different methodologies and logics. Scoring is a direct application of the Model of Hierarchical Complexity in combination with dialectical theory, choice theory and fractal theory. Micro development deconfounds micro-genetic studies, by separating task analysis, which is the basis of micro development, from performance. Rasch Analysis and the overlapping waves model are statistical analysis of performance of items.

The way the research is conducted also determines which method is to be used. Direct scoring requires some language or observed action product. Micro development requires observation of task performance and better yet a sequence of tasks at a given order, their subtasks and subsubtasks. ■

REFERENCES

Andrich, D. (2004). Controversy and the Rasch model: A characteristic of incompatible paradigms? *Medical Care*, *42*, 1-16. doi: 10.1097/01.mlr.0000103528.48582.7c

Becker, J., & Varelas, M. (2001). Piaget's early theory of the role of language in intellectual development: a comment on DeVries's account of Piaget's social theory. *Educational Researcher*, *30*(6), 22-23. doi: 10.3102/0013189X030006022

Bereiter, C. (2006). Learning and knowledge-building in open-source communities: A social-experiential approach management learning. *Management Learning, 37*(2), 187-214. doi:10.1177/1350507606063442

Boom, J., Herbert, H., & Saskia, K. (2001). Rules in the balance: classes, strategies, or rules for the Balance Scale Task? *Cognitive Development*, *16*(2), 717-735. doi:10.1016/S0885-2014(01)00056-9

Boom, J., Sanne, H. G., Van der Ven, E., Koesbergen H., & Leseman, P. P. M. (in press). Latent growth modeling and item response theory applied to the overlapping waves model.

Commons, M. L. (2008). Introduction to the model of hierarchical complexity and it relationship to postformal action. World Futures: *The Journal of General Evolution*, *64*(5-7), 305-320. doi: 10.1080/02604020802301105

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*(2), 247-267. doi:10.1080/03057240600681785

Commons, M. L., & Goodheart, E. A. (2008). Cultural progress is the result of developmental level of support. *World Futures*, *64*(5-7), 406-415. doi:10.1080/02604020802301360

Commons, M. L., Goodheart, E. A., Pekker, A., Dawson-Tunik, T. L., Cyr, E., & Rodriguez, J. A. (2005). The relationship between orders of hierarchical complexity and Rasch scaled stage scores: balance beam, laundry, and counselor-patient task sequences. *Journal of Applied Measurement. 10*(2), 1-11.

Commons, M. L., Grotzer, T. A., & Davidson, M. N. (2007). The necessity of reinforcing problem solutions for transition to formal operations: An examination of Piaget's equilibration theory of stage change. Unpublished manuscript.

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), 1-10. doi: 10.1016/0193-3973(82)90028-4

Commons, M. L., & Richards, F. A. (2002). Organizing components into combinations: How stage transition works. *Journal of Adult Development*, *9*(3), 159-177. doi: 10.1023/A:1016047925527

Commons, M. L., & Ross, S. N. (2008). The hierarchical complexity view of evolution and history. *World Futures, 64*(5-7), 399-405. doi:10.1080/02604020802301303

Commons, M., Trudeau, E., Stein, S., Richards, F., & Krause, S. R. (1998). Hierarchical complexity of tasks shows the existence of developmental stages. *Developmental Review, 18*(3), 237-278. doi:10.1006/drev.1998.0467

Dawson, T., & Heikkinen, K. (2009). Identifying within-level differences in leadership decision making. *Integral Leadership Review, 9*(5).

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, *87*(6), 477-531. doi:10.1037/0033-295X.87.6.477

Fischer, K. W., Hand, H. H., & Russell, S. (1984). The development of abstractions in adolescents and adulthood. In M. L. Commons, F.A. Richards, & C. Armon (Eds.), *Beyond formal operations: Late adolescent and adult cognitive development* (pp. 43-73). New York: Praeger.

Fischer, K.W., & Kenny, S. L.(1986). Environmental conditions for discontinuities in the development of abstractions. In R. A. Mines & K. S. Kitchener (Eds.), *Adult cognitive development: Methods and models* (pp. 57-75). New York: Praeger.

Flavell, J. H. (1963). *The developmental psychology of Jean Piaget*. Princeton, New Jersey: D. Van Nostrand Company.

Flavell, J. H. (1971). Comments on Beilin's the development of physical concepts. In T. Mischel (Ed.), *Cognitive Development and Epistemology* (pp. 121-128). New York: Academic Press.

Fodor, J. (1975). *The language of thought*. Cambridge, Massachusetts: Harvard University Press.

Fodor, J. (2008): *LOT 2: the language of thought revisited*. Oxford: Oxford University Press.

Gabriel, R. F., & Mayzner, M. S. (1963). Information "chunking" and short-term retention. *Journal of Psychology: Interdisciplinary and Applied*, *56*(1), 161-164. doi:10.1080/00223980.1963.9923710C

Gallistel, R., & Gelman, R. (2005) Mathematical Cognition. In K Holyoak & R. Morrison (Eds). *The Cambridge handbook of thinking and reasoning* (pp. 559-588). Cambridge: Cambridge University Press.

Glock-Grueneich, N. (2008). Leveraging higher education's role in social evolution: a paradigmatic strategy. *World Futures*, *64*(5), 536-553. doi: 10.1080/02604020802303986

Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C. H., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, *5*(6), 236-243. doi:10.1016/S1364-6613(00)01662-4

Herrnstein, R. J., & Vaughan, W., Jr. (1980). Melioration and behavioral allocation. In J. E. R. Staddon (Ed.), *Limits to action: The allocation of individual behavior* (pp. 143-176). NewYork: Academic Press.

Inglis, J. (2008). Evolving to address global climate change and the scale of public interactions. *World Futures*, *64*(5-7), 498-502. doi: 10.1080/02604020802303747

Inhelder, B. and Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York, NY: Basic Books.

Koplowitz, H. (2008). In praise of top-down decision making in managerial hierarchies. *World Futures*, *64*(5), 513-523. doi:10.1080/02604020802303911

Lazarsfeld, P. F. (1950). The logical and mathematical foundation of latent structure analysis. In SA Stouffer at al. (Eds) *Measurement and Prediction.* Princeton, New Jersey: Princeton University Press

Lazarsfeld, P. F. (1959). Latent structure analysis. In S. Koch (Ed), *Psychology: A study of a science, Vol 3*. New York: McGraw-Hill.

Lord, F. M., & Novic, M. R. (1968). *Statistical Theories of Mental Test Scores; with contributions by Allan Birnbaum*. Reading, Massachusetts: Addison-Wesley Publishing Co.

Miller, P. M., & Lee, S. T. (2000). *Stages and transitions in child and adult narratives about losses of attachment objects*. Paper presented at the Jean Piaget Society, Montreal, Quebec, Canada.

Miller, P. M., Lee, S. T., & Commons, M. L. (2000). *Scoring for stage and transition of losses in children and adult's lives*. Data presented at Society for Research in Adult Development Symposium, Pace University, New York City, New York.

Mislevy, R. J., &Wilson, M. (1996). Marginal maximum likelihood estimation for a psychometric model of discontinuous development. *Psychometrica*, *61*(1), 41-47. doi:10.1007/BF02296958

Miyapuram, R. S., Bapi, C. V. S. Pammi, A., Kenji D. (2006). *Hierarchical Chunking during Learning of Visuomotor Sequences*. Paper presented at International Joint Conference on Neural Networks, Vancouver, BC, Canada.

Ostini, R., & Nering, M. L. (2006). *Polytomous item response theory models*. Thousand Oaks: Sage.

Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental states. *Acta Psychologica, 32*(4), 301-345. doi:10.1016/0001-6918(70)90108-3

Pascual-Leone, J., & Goodman, D. (1979). Intelligence and Experience: A neo-Piagetian approach. *Instructional Science, 8*(4), 301-367. doi: 10.1007/BF00117011

Rasch, G. (1960/1980). *Probabilistic models for some intelligence and attainment tests*. (Copenhagen, Danish Institute for Educational Research), expanded edition (1980) with foreword and afterward by B.D. Wright. Chicago: The University of Chicago Press.

Richards, F. A., & Commons, M. L. (1990). Applying signal detection theory to measure subject sensitivity to metasystematic, systematic and lower developmental stages. In M. L. Commons, C. Armon, L. Kohlberg, F. A. Richards, T. A. Grotzer, & J. D. Sinnott (Eds.), *Adult development: Vol 2. Models and methods in the study of adolescent and adult thought* (pp. 175-188). New York: Praeger.

Robinett, T. L. (2008). A comparison of moral reasoning stages using a model of hierarchical complexity. *World Futures*, *64*(5-7), 468-479. doi: 10.1080/02604020802301618

Ross, S. N. (2008a). Fractal transition steps to fractal stages: The dynamics of evolution, II. *World Futures*, *64*(5-7), 361-374. doi: 10.1080/02604020802301196.

Ross, S. N. (2008b). Future society functioning at the paradigmatic stage? *World Futures*, *64*(5-7), 554-562. doi: 10.1080/02604020802303994

Ross, S. N., & Commons, M. L. (2008). Applying hierarchical complexity to political development. *World Futures*, *64*(5-7), 480-497. doi: 10.1080/02604020802303721

Schlaefli, A., Rest, J. R., & Thoma, S. J. (1985). Does moral education improve moral judgment? A meta-analysis of intervention studies using the defining issues test. *Review of Educational Research Fall*. *55*(3), 319-352. doi: 10.3102/00346543055003319

Scher, A. M., Young, A. C., & Meredith, W. M. (1960). Factor analysis of the electrocardiogram. *Circulation Research*, *8*(3), 519-526. doi: 10.1161/01.RES.8.3.519

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.

Spada, H., & McGraw, B. (1985). The assessment of learning effects with linear logistic test models. In S. E. Embretson (Ed.). *Test design developments in psychology and psychometrics* (pp169-194). Orlando, Florida: Academic Press.

Stein, Z., & Heikkinen, K. (2008). On operationalizing aspects of altitude: An introduction to the Lectical Assessment System for integral researchers. *Journal of Integral Theory and Practice*, *3*(1), 105-138.

Tisak, J., & Meredith, W. (1990). Descriptive and associative developmental models. In A. von Eye (Ed.), *Statistical methods in developmental research* (pp. 387-406). San Diego, California: Academic Press.

Van der Ven, S. H. G., Boom, J., Kroesbergen, E. H., & Leseman, P. M. (2011). Relations between the development of strategy selection in learning single-digit multiplication and working memory: A microgenetic study. *Journal of Experimental Child Psychology*, under revision.

William Vaughan, Jr. (1981). Melioration, matching, and maximization. *Journal of Experimental Analysis of Behavior*, *36*(2), 141-149. doi:10.1901/jeab.1981.36-141

Vaughan, W. Jr., & Herrnstein. R. J. (1987). Stability, Melioration, and Natural Selection. In Green L., and Kagel, J. H. (Eds.), *Advances in behavioral economics* (pp. 185-215). Norwood: Ablex.

Wilson, M. (1989). Saltus: A psychometric model of discontinuity in cognitive development. *Psychological Bulletin*, *105*(2), 276-289. doi:10.1037/0033-2909.105.2.276

Wright, B.D. (1992). IRT in the 1990s: which models work best? *Rasch Measurement Transactions*, *6*(1), 196-200.

Commentary on a new model for strategy development for strategy development combining categorical data analysis with growth modeling

Michael Lamport Commons

Harvard Medical School

ABSTRACT

There have been a number of steps in the evolution of modeling cognitive strategy for development. While the older stage models such as Piaget's and Kohlberg's did not have much information regarding the processes that take place between stages, Boom's new model for strategy development is 12 times as dense. It is as dense as the model of hierarchical complexity (MHC). Boom's model and MHC include substages which explain what happens between each stage of development. Existence of substages is also confirmed indirectly by Hautamaki, Marjanen, Kupiainen, and Vainikainen (2012). In the current paper, it is argued that MHC and Boom's model should be combined to have a complete model of stage development. Reasons for this proposal are discussed along with tests that can be done. Finally, few unanswered questions are posed.

KEYWORDS: cognitive strategy, strategy development, model of hierarchical complexity, sub-stages, Boom's model, latent growth modeling, item response theory, overlapping waves model

There have been a number of steps in the evolution of modeling cognitive strategy for development. First, there are staircase steps which are half stage in Kohlberg (Colby & Kohlberg, 1987), Piaget (Inhelder & Piaget, 1958), levels in Fischer (1980) and stages in Commons, Trudeau, Stein, Richards and Krause (1998). Second, there is Guttman scaling (1944, 1950) appropriate for non-probabilistic stair case models. Third steps include, Siegler's (1996) Overlapping Waves Model (OWM) which is a system (Siegler, Rest's is an earlier version of that system (Rest, 1999). Within Boom (2012), there is an implicit comparison of stages versus within stage strategies.

"Density" of a stage model

It is important to put into perspective that older stage models such as Piaget's or Kohlberg's were not very dense. They did not contain much if any explicit information filling in what happens between one stage and the next (e.g. see Colby & Kohlberg, 1987; Inhelder & Piaget, 1958). Boom's New Model for Strategy Development is 12 times as dense as Kohlberg or Piaget (Boom 2012; Commons, Commons-Miller and Miller, 2012). It has the same density as the Model of Hierarchical Complexity (MHC). MHC heavily borrows from it. Boom's model and MHC include subtasks and subsubtask which fill in information about what happens between one stage and the next which older stage models such as Piaget's or Kohlberg's did not. For example, in primary stage of arithmetic there are three subtask actions, each containing five subsubtask actions.

Subsubtasks are necessary for acquisition of subtask action (Boom, 2012). However, subsubtasks do not persist after acquisition whereas subtasks persist even upon completion of the next stage.

More indirect confirmation of sub-stages

Hautamäki, et. al. (2012) also gave indirect confirmation of the existence of subtasks. They do so by showing that different comparisons within the Water Level Task (WLT) vary in difficulty. It is important to note that this work is a very unusual task for a human to do. It is one that does not appear to have a great deal of evolutionary significance. The WLT is like the formal balance beam task of the MHC, but even one stage more difficult. It combines understanding volume and the tilt level. That is, motorically it is easily solved, but understanding how it works is much more difficult (probably systematic order).

» WHAT IS NEEDED TO HAVE A COMPLETE MODEL OF STAGE DEVELOPMENT?

We argue that the model of hierarchical complexity is a metasystem. Boom's new model for strategy development is a metasystem as well. They need to be combined.

Boom's model is a metasystem stage 12

Boom's model is a metasystem stage 12 because it combines two systems and applies them to the OWM. Latent growth model (LGM) is the first system and item response theory (IRT) is a second system. The IRT provides the means to relate the use of such strategies to an underlying developmental dimension and the movement

of individuals along this dimension can be modeled using LGM (Boom, in press). To combine the systems, Rasch latent scores for group data is required. Learning how to characterize individual scores as to slope and intercept to be compared is required as well. All these are combined in Boom's model which fully coordinates the systems yielding a metasystem.

The transition to the paradigmatic stage 13: Why using both Boom's new model for strategic development and an analysis of the tasks using MHC is necessary

Neither of the two metasystems is sufficient alone because psychometric systems do not have a priori predictive systems. Without Boom's model we cannot check those predictions. What is needed is to combine two metasystems: Boom's new model for strategic development and Common's three layer model of hierarchical complexity. The one of Boom is a psychometric model of performance. This would also apply to Hautamäki. Common's three layer model of hierarchical complexity is one of orders of hierarchical complexity, it is one of subtasks required actions and sub-subtasks required actions.

The MHC is a metasystem. Boom's model is a metasystem. Combining two metasystems is paradigmatic. However, because the combination has not been tested predictively, it is the last subtask in the transition to paradigmatic stage.

Towards a psychophysics of development

On the stimulus side there are required subsubtask actions. This is what is captured by the MHC. These required sub-subtask actions are task-stimulus for the strategies which are the behaviors. This is what is captured by the OWM. When the two models, the MHC and the psychometrics, combine they produce the new paradigm which is the psychophysical paradigm of development.

Tests

One proper test is showing that the priori subtasks are complete and predict the strategies between stages and the orderings of the stages themselves. The latter has been done with the stages showing that there are gaps between stages and that they are equally spaced (Commons et al., 2014). Likewise, we might predict that the spacing between subtasks within an order are equally spaced. This can be tested using Boom's model. The subtask numbers and a combination index combining order tasks and subtasks, as well as within stage strategies across many stages can be used to predict stage strategies.

Unanswered questions

With distinctions between stage strategies on one hand and micro developmental strategies between stages on the other, there may be some confusion. In the MHC there are 17 stages. Within each stage, are there always the same number of subtasks? Do they vary with task sequence, domain and order? Does the OVM help us identify missing subtask and sub-subtask actions? Does it help us identify superfluous subtask and sub-subtask actions? Can it help us distinguish between sufficient and necessary sub-subtask actions? What would be the expected r's between sub-subtask action number, order number and Rasch score? ■

REFERENCES

Boom, J. (2012, July). *New Model for Strategy Development: Combining Categorical Data Analysis with Growth Modeling*. Paper presented at Biennial Meeting of the International Society for the Study of Behavioral Development, Edmonton, Canada.

Colby, A. & Kohlberg, L. (1987). *The measurement of moral judgment*. NY: Cambridge University Press.

Commons M. L., Commons-Miller L. A. H., & Miller P. M. (2012, July). *Boom's Overlapping Wave Model, Rasch Analysis and Subtasks in Transition Between Primary and Concrete Stage Arithmetic.* In N. Galambos (Chair), Quantitative Analysis of Qualitative Developmental Progress: Modeling Cognitive Strategy Development. Symposium conducted at the meeting of International Society for the Study of Behavioral Development, Edmonton, Alberta, Canada

Commons M. L., Li, E. Y., Richardson, A. M, 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).

Commons, M. L., Trudeau, E. J., Stein, S. A., Richards, F. A., & Krause, S. R. (1998). The existence of developmental stages as shown by the hierarchical complexity of tasks. *Developmental Review, 8*(3), 237-278. doi: 10.1006/drev.1998.0467

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, *87*(6), 477-531. doi:10.1037/0033-295X.87.6.477

Guttman, L.A. (1944). A basis for scaling qualitative data. *American Sociological Review, 91*, 139-150.

Guttman, L.A. (1950). *The basis for scalogram analysis*. In Stouffer, S.A., Guttman, L.A., & Schuman, E.A., Measurement and prediction. Volume 4 of Studies in social psychology in world war II. Princeton: Princeton University Press

Hautamaki, J., Marjanen, J., Kupiainen, S., & Vainikainen, M.P. (2012, July). *Old wine in a new bottle - Modelling the development of co-ordination of horizontal and vertical planes using WLT-task*. In N. Galambos (Chair), Quantitative Analysis of Qualitative Developmental Progress: Modeling Cognitive Strategy Development. Symposium conducted at the meeting of International Society for the Study of Behavioral Development, Edmonton, Alberta, Canada

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

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York, NY US: Oxford University Press.

Rest, J.R., Narvaez, D., Bebeau, M. & Thoma, S. (1999). *Postconventional Moral Thinking: A Neo-Kohlbergian Approach*. Mahweh, NJ: Lawrence Erlbaum Associates.

Van der Ven, S. H. G., Boom, J., Kroesbergen, E. H., & Leseman, P. M. (2011). Relations between the development of strategy selection in learning single-digit multiplication and working memory: A microgenetic study. *Journal of Experimental Child Psychology*, under revision.

Van der Ven, S.H.G., Boom, J., Kroesbergen, E., & Leseman P.P.M. (2012). Microgenetic patterns of children's multiplication learning: Confirming the Overlapping Waves model by latent growth modeling. *Journal of Experimental Child Psychology*,*113*(1), 1-19. doi:10.1016/j.jecp.2012.02.001

Correspondence between some life-span, stage theory developmental sequences of stages and levels

Charu Tara Tuladhar[1] and Michael Lamport Commons[2]

1 Mount Holyoke College
2 Harvard Medical School

ABSTRACT

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.

KEYWORDS: developmental stages, sequence, model of hierarchical complexity, moral judgment, Kohlberg, Piaget, moral maturity scores, Fischer, conversion, cognitive

There are many good comparisons of developmental sequences from the perspective of different theories. However, they do not cover an entire life-span. Commons, Trudeau, Stein, Richards, & Krause (1998) constructed a table showing the relationship among stage models including the models of Commons, Richards and Armon, 1984; Commons et al. (1998); Fischer and Bidell (1998), Colby and Kohlberg (1987a, 1987b); Case (1985), Campbell and Bickhard (1986); and Piaget and Inhelder (1969). More recently, Dawson-Tunik, Commons, Wilson and Fischer (2005) constructed a comparison table of development sequences comparing developmental stages of Dawson-Tunik (2004), Piaget and Inhelder (1969), Fischer and Bidell (1998), Commons et al. (1998); Colby and Kohlberg's (1987a, 1987b) 9 point scale; Armon (1984); and King and Kitchener (1994). Among these models, Colby and Kohlberg's (1987a, 1987b) 9 point and 13 point scales have been widely used (e.g. Kegan, 2002) as Kohlberg's theory expands on Piaget's work. Kohlberg determined that the process of moral development was primarily concerned with justice, and claimed that it continued throughout an individual's lifetime (Kohlberg, 1981). Many comparisons have been made among various stages of development previously but, researchers have only used Colby and Kohlberg's 9 point scale for the comparisons. There has not been any work that shows how Colby and Kohlberg's 13 point scale translates into other developmental sequences. Kohlberg's

model is limited to assessing the development of moral judgment only. The model of hierarchical complexity (MHC) (Commons et al. 1998), on the other hand, is a general stage model that assesses development in any domain. The model helps score how hierarchically complex a behavior is (Commons, Trudeau, et al. 1998). It has 17 developmental stages. This paper shows the correspondence among orders of hierarchical complexity (OHC) and the 13 point scale of moral judgment, the corresponding 9 point scale, Fischer and Bidell's cognitive development level and Piaget and Inhelder's cognitive development stages. These five models of development were chosen because the stages in these models cover the developmental processes that occur in an entire life-span of an individual. Other developmental models were excluded as they do not cover the entire life-span. There have been several tests that validate these models. Here, we construct such a comparison that goes lower and higher than those in the literature cited above.

» THE CORRESPONDENCE TABLE

Table 1 presents the stages of the model of hierarchical complexity (MHC) and the corresponding stages of Fischer and Bidell (1998), Piaget and Inhelder (1969), Colby and Kohlberg's (1987a, 1987b) 9 point scale and 13 point scale and their respective moral maturity scores (MMS). MMS was described by Colby et al. (1983) as a measure of the moral judgment stages. The score is a continuous variable representing the proportion of moral reasoning done by individuals at each stage of Kohlberg multiplied by the ordinal

Correspondence regarding this article should be addressed to Dr. Michael Lamport Commons, Harvard Medical School, 234 Huron Avenue, Cambridge MA 02138. E-mail: commons@tiac.net

number of that stage. For example, an MMS of 200 indicates that all of the individual's reasoning is at stage 2 of the 9 and 13 point scales of moral judgment and an MMS of 300 indicates that all of the individual's reasoning is at stage 3 of the 9 and 13 point scales of moral judgment. However, in the current paper, the distribution of the MMS to the stages of Kohlberg and descendants has been adjusted. The stages of Fischer and Bidell, and Piaget & Inhelder that correspond to the orders of hierarchical complexity were adapted from the conversion tables provided by Commons, Trudeau, Stein, Richards & Krause (1998) and Dawson-Tunik, Commons, Wilson, & Fischer, (2005). The conversion of Kohlberg and decedents' 9 point scale of moral judgment and 13 point scale of moral judgment into the orders of hierarchical complexity (OHC) was made on the basis of the following three assumptions.

1. Model of hierarchical complexity is model that measures development and shows sequence of actions for a task. There has been a lot of empirical evidence that substantiates not only the face validity of this model, but also the extremely high predictions of Rasch scaled performance from the orders of hierarchal complexity of tasks—up to $r = .984$ (e.g. Commons et al., 2014, Commons et al., 2008, Commons et al., 2006).

2. 50 point rule: The moral maturity scores (MMS) of each order of hierarchical complexity are 50 scores apart. For example, an MMS of 100 indicates that an individual is performing at stage 5. An MMS of 150 indicates that an individual is performing at stage 6. An MMS of 200 indicates that an individual is performing at stage 7 and so on. Hence, as orders of hierarchical complexity increase by 1, the corresponding MMS score increases by 50 points. As Pascual-Leone (1972) showed, all the half stages of Piaget and therefore of Kohlberg are really full stages. Thus, each half stage of Kohlberg would be 50 MMS apart.

3. The model of hierarchical complexity (MHC) applies to Inhelder and Piagetian (1958) theory of stage that two or more lower order actions constitutes one action of a higher order of complexity. Those actions have to be coordinated. However, MHC also adds that the ordering of the lower order actions should be non-arbitrary.

The conversion of Kohlberg and descendants' 9 point scale of moral judgment to the orders of hierarchical complexity was derived by scoring the definitions used in Kohlberg's moral judgment instrument. If it was not absolutely clear, the example from the Colby and Kohlberg manual was used. The equation was:

$3 + 2 \times$ *(stage number of 9 point Colby and Kohlberg)* = OHC

Here, 2 and 3 are constant numbers. The constant 3 aligns the OHC and Colby and Kohlberg stage. Multiplying by 2 converts the half stage numbers of Kohlberg stages that were really full stages into full number.

There were only a few major changes made on the 9 point scale. On the high end, stage 4/5 was scrapped from Kohlberg's stages because stage 4 of Kohlberg corresponds to Systematic stage (Stage 12) of MHC and the transition to stage 5 of Kohlberg corresponds to metasystematic stage (stage 13) of MHC. According to the 50 point rule for MHC, stages 12 and 13 of MHC are supposed to be 50 MMS apart which means that stages 4 and 5 of Kohlberg would also have to be 50 MMS apart. However, according to Kohlberg,

stages 4 and 5 of the Kohlberg stages are 100 MMS apart. Also, following the 50 point rule, the half stage, 4/5, of Kohlberg's had to be scrapped. Sonnert and Commons (1994) found that Stage 5 and 6 were actually part of a single stage and were consolidated into a single stage 6 which has an MMS of 500. Thus, new moral maturity scores were assigned to stages 5 and 6 of the 9 point scale to preserve the consistency of the 50 point rule. For example, according to Kohlberg and his descendants, stage 5 is assigned 500 MMS, stage 6 is assigned 600 MMS and so on. However, after our adjustments, stage 5 of the 9 point scale was assigned 450 points, stage 6 was assigned 500 points and stage 7 was assigned 550 points.

Three higher stages were introduced, including most importantly the paradigmatic stage 14 which is stage 6 in Colby and Kohlberg's stages of moral judgment (Sonnert & Commons, 1994). Kohlberg's speculative stage 7 did not meet any of the stage considerations required for hard stages or for the MHC. Thus, it was rejected. The stage 7 of Colby and Kohlberg in Table 1 was an added stage which corresponds to the cross-paradigmatic stage 15 of OHC. Similar adjustments were made to the 13 point scale. On the low end of the scale, below stage 1 (Colby & Kohlberg, 1987), almost everything had to be redone. We again applied the 50 point rule to the corresponding orders of hierarchical complexity. Colby & Kohlberg's stages begin at stage 1. Stage 1 of Colby and Kohlberg corresponds to stage 5 of the MHC. Thus, we extended the Kohlberg stages down to stage (-1/-2) using the 50 point rule of MMS.

The 9 point scale and the 13 point scale refer to the same stages of moral judgment. They differ only in the way their substages were divided. For example, on the 9 point scale, the transitional stage between 2 and 3 is 2/3 whereas in the 13 point scale, the transitional stages between 2 and 3 are 2(3) and 3(2). The stages on the 9 point scale are divided by half whereas the stages in the 13 point scale are divided by one third. Thus, the conversion of 13 point scale of moral judgment to the stages of model of hierarchical complexity was induced by following the 50 point rule.

It is also important to note that there are stages in the Colby and Kohlberg's 13 point scale that do not correspond to the MHC stages and are between the MHC stages (e.g. 1(2) between MHC stages 6 and 7; 2(1) between MHC stages 7 and 8). These 13 point scale stages are not really stages, but could be possible transitional from one stage to the next.

» CONCLUSION

In this paper, a correspondence table that compares five life-span developmental sequences was presented. In addition, their corresponding moral maturity scores were also given. The table included the orders of hierarchical complexity, Fischer and Bidell's (1998) stages of cognitive development, Piaget and Inhelder's (1969) stages of cognitive development, Colby and Kohlberg's (1987a, 1987b) 9 point scale of moral judgment and Colby and Kohlberg's (1987a, 1987b) 13 point scale of moral judgment. Adjustments were made to Colby and Kohlberg's stages. The 13 point scale of Colby and Kohlberg was presented in a correspondence table for the first time. This table allows one to intelligently use Kohlberg's scoring manual and easily see how stages of different stage models of developmental sequences correspond to each other. ∎

Table 1. General description of sequence

Charu Tara Tuladhar
Michael Lamport Commons
September 6, 2012

Discriminations	Order of hierarchical complexity Commons et al. (1998)	Fischer & Bidell's stages of cognitive development (1998)	Piaget & Inhelder's stages of cognitive development (1969)	9 point scale of moral judgment Colby and Kohlberg (1987)	13 point scale of moral judgment Colby and Kohlberg (1987)	Moral maturity scores
Calculatory	0					
Automatic	1	-1**		-1/-2**		-150
					-1(-2)**	-133.33
Sensory or motor and not both	2	0		-1**	-1**	-100
					-1(0)**	-66.66
Circular sensory motor	3	1	a Sensorimotor	0/-1**		-50
					0(-1)**	-33.33
Sensory motor	4	2	b Sensorimotor	0**	0**	0
					0(1)**	33.33
Nominal	5	3	Ia Preoperational	0/1**		50
					1(0)**	66.66
Sentential	6	3-4*		1	1	100
					1(2)	133.33
Preoperational	7	4	Ib Preoperational	1/2		150
					2(1)	166.66
Primary	8	5	IIa Preoperational	2	2	200
					2(3)	233.33
Concrete	9	6	IIb Concrete operational	2/3		250
					3(2)	266.66
Abstract	10	7	IIIa Concrete operational	3	3	300
					3(4)	333.33
Formal	11	8	IIIb Formal operational	3/4		350
					4(3)	366.66
Systematic	12	9	IIIc Formal operational	4	4	400
					4(5)*	433.33
Metasystematic	13	10	Postformal**	5*	5*	450
					5(6)*	466.66
Paradigmatic	14	11*	Postformal**	6*	6*	500
					6(7)*	533.33
Cross-paradigmatic	15	12*	Postformal**	7**	7**	550
					7(8)**	566.66
Meta-paradigmatic	16	Non-existent	Not observed			600

Notes: *Speculated stages that correspond to the orders of hierarchical complexity. ** Stages that do not exist in the models but added here based on how they would correspond to the orders of hierarchical complexity.

REFERENCES

Armon, C. (1984). Ideals of the good life and moral judgment: Ethical reasoning across the life span. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), *Beyond formal operations: Vol. 1. Late adolescent and adult cognitive development* (pp. 357-380). New York: Praeger.

Campbell, R. L., & Bickhard, M. H. (1986). *Knowing levels and developmental stages*. Basel; New York: Karger

Case, R. (1985). *Intellectual development: Birth to adulthood*. Orlando, Florida: Academic Press

Colby, A., Kohlberg, L., Gibbs, J., Candee, D., Speicher-Dubin, B., Kauffman, K., Hewer, R., & Power, C. (1983). *Measurement of Moral Judgment: A Manual and its Results*. New York: Cambridge

University Press

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. (Eds.) (1987b). *The measurement of moral judgment: Vol. 2. Standard form scoring manuals. New* York: Cambridge University Press.

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. 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.

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).

Commons, M. L., Richards, F. A., & Armon, C. (Eds.). (1984). *Beyond formal operations: Vol. 1. Late adolescent and adult cognitive development.* NY: Praeger.

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*(6), 430-435.

Commons, M. L., Trudeau, E. J., Stein, S. A., Richards, F. A., & Krause, S. R. (1998). Hierarchical Complexity of tasks shows the existence of developmental stages, *Developmental Review,18*, 237–278.

Dawson-Tunik, T. L. (2004, November). The Lectical TM Assessment System. 1. Retrieved December, 2004, from http://www.lectica.info.

Dawson-Tunik, T.L., Commons, M.L., Wilson, M. & Fischer, K.W. (2005). The shape of development, *European Journal of Developmental Psychology*, *2*(2), 163–195.

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

Kegan, R. (2002). A guide to the subject-object interview. Unpublished Scoring manual. Harvard Graduate School of Education.

King, P. M. & Kitchener, K. S. (1994). *The development of Reflective Judgment in adolescence and adulthood* . Jossey Bass: San Francisco.

Kohlberg, L. (1981). *Essays on Moral Development, Vol. I: The Philosophy of Moral Development*. San Francisco, CA: Harper & Row.

Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York, NY: Basic Books.

Sonnert, G., & Commons, M. L. (1994). Society and the highest stages of moral development. *Politics and the Individual*, *4*(1), 31-55.

Fractal model of nonlinear hierarchical complexity: measuring transition dynamics as fractals of themselves

Sara Nora Ross

John F. Kennedy University
Antioch University Midwest

ABSTRACT

Fractal transition theory and measurement enable fine-grained analysis of the most seemingly-chaotic of the developmental transition phases. The explication of the fractal nature of those transition dynamics informs study of learning, decision making, and complex systems in general. A hallmark of the fractal measure is the use of thesis-organized transition measures that are orthogonal to time. Using this method, unpredictable behaviors become "rational" when understood in terms of attractors within developmental processes. An implication for nonlinear science is to transform data otherwise interpreted as incoherent "white noise" into the coherent fractals of the "pink noise" dimension. By integrating Commons et al's Model of Hierarchical Complexity (MHC) and this nonlinear model of the fractal transitional orders of hierarchical complexity, a unified mathematical theory of behavioral development will be possible. Such a new formal theory would account for the entire span of behavioral development's equilibrium states and phase transitions, from lowest to highest orders of complexity. The mathematical expressions for the transitional orders of hierarchical complexity must be developed and integrated with the existing MHC.

KEYWORDS: behavioral development, decision making, developmental transitions, fractal model of nonlinear hierarchical complexity, fractal transition theory, learning, microdevelopment, nonlinear dynamics, phase transitions

This paper builds on my previous work (Ross, 2007, 2008; Ross & Commons, 2007) to introduce fractal transition theory and measurement, and on that basis to propose the fractal model of nonlinear hierarchical complexity, and to discuss related issues. One intended contribution is to add to the methodological discourse on transitions in the developmental field, and another is to help forge closer linkages between that field and complexity science. A significant motivation for the paper is to attract expertise to help with the mathematical expression of the transitional orders of complexity presented herein.

To accomplish these purposes, the paper is organized as follows. After situating this contribution within the literature on dynamic transitions, the body of the paper is a critical discussion of pre-fractal and fractal transition measurements used to date in conjunction with the model of hierarchical complexity (see Commons, in this issue). My focus is on the most chaotic phase of transition processes. Within that discussion, I supplement my original description of the fractal nature of that transition and explicitly propose the new model. The closing discussion offers key implications of fractal insights into developmental transitions for behavioral sciences.

» TRANSITION DYNAMICS: UNDERSTANDINGS AND METHODS

Transition refers to one or more movements of a system from one state, phase, or activity to another. As a property of dynamics systems, transitions are widely studied. Transition step sequences between stages of performance situate the model of hierarchical complexity (MHC; Commons, Goodheart, Pekker, et al. 2007; Commons, Trudeau, et al. 1998) in the specialized field of microdevelopment as well as the neo-Piagetian tradition. Its content-free and scale-free orders of complexity are fractal by definition (they apply to any actions), properties that situate the MHC in the complexity sciences that study nonlinear systems' phase transitions. Microdevelopment studies include dynamic systems approaches to study humans as developing dynamic systems; the field recognizes transitions in task completion occurring over time scales from minutes to months (Granott & Parziale, 2002). They can also take as long as years for highly complex endeavors (Commons, Ross, & Bresette, 2011; Fischer & Yan, 2002). These and other fields of study pay close attention to transitions over time, because from transitions, new behaviors are constructed and emerge. Predictably, a diverse range of methods are used to study and measure transitions.
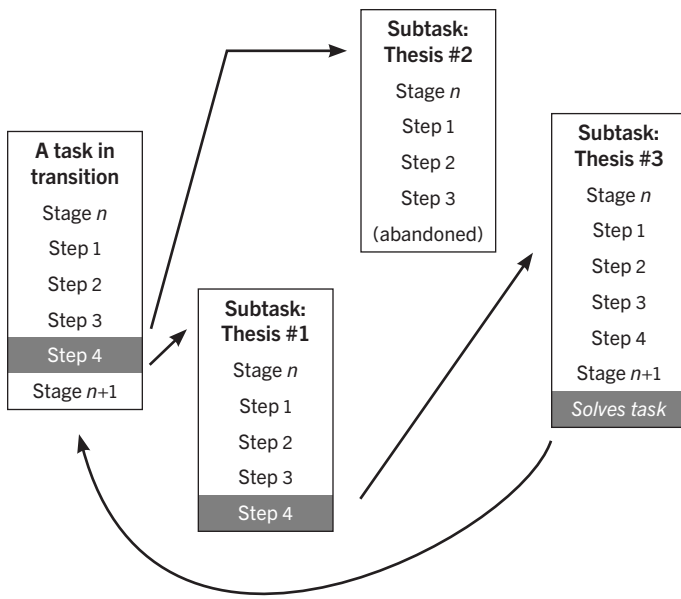
Figure 1. Representation of smash phase fractals of transitional orders of hierarchical complexity

Methods in nonlinear sciences use mathematics and graphic data analysis to study phase transitions, and transitions have familiar names like bifurcations, catastrophes, and oscillations to distinguish them; some are more complex dynamics than others. Yet nonlinear methods have not yet recognized vertical increases in complexity, i.e., the developmental implications of many transitions. This is the case despite the ubiquitous discussion of self-organization and emergence, and widespread recognition of chaos preceding self-organization. Goldstein (2002, 2007), however, began making inroads to develop a theory of emergence with increasing levels of complexity, using the concept of self-transcending constructions.

Microdevelopment studies, with their focus on learning and development, may use dynamic systems methods to examine developmental transitions of specific tasks (e.g., infant motor learning, Thelen & Smith, 1994), narrative methods to describe generalizable patterns (e.g., Kuhn, 2002), and ordinal scales or coding schemes (e.g., Basseches & Mascolo, 2010; Gelman, Romo, & Francis, 2002; Parziale, 2002).

Sabelli (1995, n.p.) proposed an interdisciplinary bridge to "interpret non-linear dynamics as a mathematical formulation of dialectic logic" that represent a phase plane in four quadrants of dialectical conditions familiar to Piagetians. The dialectic of thesis, antithesis, synthesis is commonly distilled as *A; B (or not A); A or B; A with B*. Such neo-Piagetian orientations emphasize the repeating patterns of transitions from less to more complex stages. That dialectical tradition is explicit in transition work done by Yan and Fischer (2007), Basseches (1984), Laske (2009), Commons and Richards (2002), Basseches and Mascolo (2010), and Ross (2007, 2008; Ross & Commons, 2007). Yan and Fischer (2007, p. 59) summarized the dynamics in individuals' learning as "change among these patterns unstable, fluctuating, stable in a continuous process of self-organization that produces the four types of trends—disorganization, regression, improvement, and stabilization—across sessions" and related at least some aspect of the performances to developmental skill theory levels.

Beyond the dialectical schemes, greater numbers of transition dynamics between stages have been discriminated using narrative and coding methods. Laske's (2009) analyses of interviews incorporated the narrative dialectical transition schemata developed by Basseches (1984), 24 progressions of content-types or "thought forms" that construct dialectical thought. When they reported the addition of transition steps to the model of hierarchical complexity, Commons & Richards (2002) renumbered and added to Piaget's original steps. They based new substeps on premises of choice and signal detection theories, "based on Kuhn and Brannock (1977) and the systematization of that by Commons and Richards (1984b)" (Commons & Richards, 2002, p. 162). Richards and Commons (1990) had proposed using that same signal detection approach in structured experiments to test for existence of some higher stages.

A prominent commonality across the methods reviewed thus far is the linear time dependence of the measures and thus the analyses. There seems to be only one exception to this norm: the use of theme-organized transition scoring that is orthogonal to time; that is, measurement that is not tethered to the time axis. This means the task scoring remains associated with time but is not organized or measured by linear timing of the dynamics. Both the coactive systems coding process for two-person systems developed by Basseches and Mascolo (2010) and the fractal transitions approach developed by Ross (2007, 2008; Ross & Commons, 2007) enabled fine-grained moment-to-moment analysis of transitions that continue over time, revealing interactional dynamics invisible if chronology drove the analysis. Independently, these researchers found that the true nonlinearity of human behavior is perhaps best revealed by using methods that track the structure and process of changes, regardless of their time stamp. Central to these methods is to identify each thesis-action that emerged and track every associated action until an eventual synthesis completes the transition or the task is abandoned. Basseches and Mascolo's coding scheme is not designed to surface fractal patterns. Thus far, it seems fractal methods for developmental transition analysis appear in only my work cited above. In this paper I supplement my original description of the fractals in the "chaotic" smash phase transition to explicitly propose new constructs of transitional orders of hierarchical complexity as part of my fractal transition theory.

» STAGE TRANSITION MEASUREMENTS USED WITH MODEL OF HIERARCHICAL COMPLEXITY: PRE-FRACTAL METHODS

Hierarchical complexity scoring approaches (Commons, Rodriguez, Miller, Ross, LoCicero, et al, 2007) adopted the expanded transition step scheme proposed by Commons and Richards (2002) to measure transitions from one stage of performance (*n*) to another (*n+1*). That scheme recognized two more kinds of phase-shift dynamics occurring within Piaget's dialectical sequence. The basic ordinal scoring is summarized as follows (adapted from Commons & Richards, 2002, p. 162).

Stage n: Entity operates with temporary equilibrium *A* (thesis) until transition begins at step 1.

Table 1. Patterns of coherence in data measured with fractal method

Orders of Hierarchical Complexity	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
12-2																	5																					
12-1																5																						
12-0																																						
11-4																											7											
11-3										3																7												
11-2									3																7													9
11-1								3																7													9	
11-0							3					1										6	7													9		
10-4											1								4		2							2							9			
10-3			1			2									4			4		2														9				
10-2		1			2									4																	8		9					
10-1	1			2									4																	8		9						
10-0																													8									
Duration in secs	18	18	22	17	1	14	5	3	12	5	45	5	14	19	5	33	14	27	57	12	13	30	6	19	2	14	17	22	3	4	10	3	1	14	42	35	9	24
Item #	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
Item Scores	10-1	10-2	10-3	10-1	10-2	10-3	11-0	11-1	11-2	11-3	10-4	11-0	10-1	10-2	10-3	12-1	12-2	10-3	10-4	10-3	10-4	11-0	11-0	11-1	11-2	11-3	11-4	10-4	10-0	10-1	10-2	10-1	10-2	10-3	10-4	11-0	11-1	11-2

Sequential Actions over 11-Minute Trial

Step 1: *A* not true—destabilization

Step 2: *B* (or not *A*)—negation (antithesis)

Step 3: *A or B*—oscillation (relativism)

Step 4: *A and B*—"Chaos" (attempts at synthesis: smash)

Stage n+1 A with B forms new action *C*—New temporary equilibrium *C* (synthesis)

Using this scoring system one can quantify the occurrence and progression of transition processes in task performances at any Order of Hierarchical Complexity (OHC). It affords meaningful explanatory power for the "how" of development from one stage to another, in any domain of task actions. It is *coarse-grained* with respect to the step 4 smash dynamics because one cannot use it to score the tasks performed during attempts to reach synthesis.

Therefore, the focus in this paper is on explicating the complex dynamics within step 4's "smash" transition phase. That focus is driven by my interests (*a*) to enable use of nonlinear methods to measure developmental behavioral dynamics at any scale, and (*b*) to complete the foundations for a *fractal model of nonlinear hierarchical complexity*. My interest in smash, however, was preceded by others' work. Their early focus on explicating smash used signal detection and choice theory methods for empirical purposes, as mentioned earlier.

> *The following steps are not stages in the sense of the general stage model subsequently named the model of hierarchical complexity. They are not analytical constructs having the necessary properties of orders and hierarchical complexity and the resulting stages. The steps belong to the realm of empirical science and describe the steps of stage acquisition in an empirically testable manner. (Commons & Richards, 1995, p. 7)*

The three substeps were developed to "describe different ways of smashing *A* and *B* together, without fully coordinating them…

1. Smash$_1$ hits and excess false alarms and misses

2. Smash$_2$ hits and excess false alarms

3. Smash$_3$ correct rejections and excess misses" (Commons & Richards, 2002, pp. 162–163)

As constructs for experiments with predetermined options and answers, these substeps are *nominal:* they represent *sets of categories* of certain actions, and those actions are nominally described in metaphorical terms. In short, this substep scheme does not support mathematical expression of single actions. It has to be confined to experimental settings that can use it and it has to be excluded from hierarchical complexity transition theory and measurement. I learned the hard way that these substeps are also a mismatch for scoring natural behaviors of an entity (person, group, system, etc.). In all natural behaviors, each person or group generates a unique set of variables to coordinate in the process of eventually arriving at its own synthesis: no one can predict which variables will emerge, or be rejected, or be incorporated in a synthesis until it happens!

For general use in developmental analysis, the straightforward four-step scoring scheme above is sufficient; few of us have a need to delve into micro analysis of the smash step. Yet, in the course of my efforts to measure the nonlinear task dynamics of smash

phase transitions, I discovered the behavioral fractals that comprise them (Ross, 2007), and their theoretical and scientific implications (Ross, 2008). The fine-grained fractal transition measurement approach is presented next.

» FRACTAL METHOD

The significance of presenting the fractal transition measurement is directly connected to the formal, general theory status of the model of hierarchical complexity (MHC; Commons, Goodheart, Pekker, et al. 2007). As a mathematically-based, universal, scale-free behavioral development model, the MHC accounts for the discrete orders of hierarchical complexity of actions. Those orders have enabled us to measure stages of developmental performance for over 30 years. Currently, MHC explains *that* discrete orders of hierarchical complexity are constructed by coordinating lower order actions, but does not yet explain the *how* of those coordinations. Those coordinating actions are discrete for measurement purposes yet occur during continuous living system behaviors. Thus, the MHC does not yet describe *transitional* orders of complexity. Consequently, no formal theory yet accounts for the *continuity* of actions' emergence comprising behavioral development.

To possess "universal, scale-free" properties means the MHC's orders of hierarchical complexity are fractal. Fractal means the repetition of self-similar patterns at different scales. Behavioral scales from the micro-biological to large social systems evidence the orders of hierarchical complexity (see Commons & Ross, 2008). The fractal transition theory is proposed as a universal, scale-free general model as well (Ross, 2008). Its measurement is discussed next.

On the surface step 4's A *and* B implies only two actions are involved in smash. As the substeps from Commons and Richards suggested, many actions may be performed during this phase. The fractal transition measure accommodates the unpredictable origination of new actions as an entity constructs them nonlinearly in real time. In individuals and groups, these ubiquitous dynamics appear in behavioral processes of learning, reflecting, explaining, problem identification, problem solving, decision making, meaning making, perspective taking, theorizing, and so on.

The transitional orders of hierarchical complexity are ordinal, consistent with the MHC. They are expressed in terms of MHC *primary order's ordinal* paired with the *transition order's ordinal*. For example, a formal operations task performance in the oscillation phase (A *or* B) is indicated as 10–3 for primary order 10, transitional order 3. Every transition begins with some temporary equilibrium (A), regardless of when or where a transition occurs. This means transitions nested within transitions are ordered in exactly the same way.

The fractal measurement uses the same math-based transition steps 1 through 3 as presented earlier. If transition is not abandoned after step 3, the smash phase begins with the entity's next action, which is to construct a thesis related to the original task. It is another A, temporary equilibrium. What happens next is unpredictable. The new thesis may launch one or more subtasks with varying degrees of nesting within other distinct theses. Regardless of when they emerge in the process, smash subtasks measure as full or partial fractals of the transitional orders of hierarchical complexity (Figure 1).

Once constructed, smash-phase theses are commonly developed as well as temporarily abandoned in a discontinuous fashion. The discontinuity is not a measurement problem. By their nature, the fractal measures transform seemingly random data to reveal their observably coherent order. This is because the fractal transitional orders are orthogonal to time, and transition sequences begin with measureable theses. Figure 2 displays data to illustrate these points.

» The 38 items were sequentially-spoken actions during an 11-minute decision-making (problem-solving) session, and scored using the fractal method. The 38 item numbers are listed along the time axis at the bottom of the figure.
» Items' duration in seconds are shown along the time axis.
» The scores of the sequentially-spoken actions are on the horizontal time axis under their respective item numbers.
» The vertical axis is the relevant range of ascending transitional orders of hierarchical complexity. Actions are measured against them.
» The filled cells communicate two kinds of information.
  » The number in a cell references one of the nine theses constructed during the trial. Thesis numbers range from 1-9.
  » The *placement* of the thesis-number in a cell indicates where an action fell on the ordinal scale of complexity. That placement yields the scores given on the horizontal time axis.

If the chronologically sequenced item scores in the bottom row of Figure 2 were plotted on a graph, their erratic discontinuity would be starkly obvious. By contrast, when the data associations with each thesis are maintained, clearly coherent patterns are evident. The patterned data indicate *how* the entity is developing more complex behaviors from moment to moment, and demystify *how* syntheses at higher order task performances are constructed.

» DISCUSSION

Living entities are dynamic systems that behave nonlinearly and thus unpredictably, yet always with coherence when we use measures that "let their data speak." When they do, we can "hear" them without distortions of linear time-based assumptions and methods. With measures based on the fractal transitional orders of hierarchical complexity, we can "give voice" to data generated by nonlinear behaviors via the fine-grained analysis of the most complex transition phases.

These unpredictable behaviors are "rational" when understood in terms of *attractors* operating within developmental processes. Each thesis constructed by an entity is an attractor, and the coordination processes return to it at different points in time until it is resolved, or cannot be resolved and is abandoned, or is interrupted and subsequently forgotten or conditions change. This accounts for why there are often-discontinuous actions on a thesis that are nonetheless developmentally coherent. As I previously argued in more detail (Ross, 2008), there is an important implication here for nonlinear science methods: the transformation of data otherwise interpreted as incoherent "white noise" into the coherent fractals of the "pink noise" dimension. Further—and for the first time, I believe— with hierarchical complexity transition measures, complexity science could discriminate if/how systems' phase transitions result in hierarchically greater system complexity. The meaning and implications of many transitions could become more evident and more deeply understood.

Users of the model of hierarchical complexity's previous scoring system now have direction on a scoring method that is internally consistent with the general model. The earlier non-theoretical substeps have theoretically-sound replacements for those who do fine-grained analysis.

I hope one implication of this work for adult development specialists is reinforcement of the understanding that behavior develops in any domain task by task: people are not "at" a stage of development, but rather, day in and day out, they perform tasks at different stages of development and much of the time tasks are in transition phases.

Finally, this work implies that a unified theory of behavioral development is on the horizon: Commons et *al.'s* model of hierarchical complexity and this nonlinear model of the fractal transitional orders of hierarchical complexity demand integration. The resulting formal theory would account for the entire span of behavioral development's equilibrium states (satisfied by the current MHC) as well as the phase transitions (this current proposal) from lowest to highest complexity. The mathematical expressions for the transitional orders of hierarchical complexity must be developed. A large *n* study will be vital so the contributions of these nonlinear developmental measures find their way into sciences of learning, decision making, and complex systems. ■

REFERENCES

Basseches, M. (1984). *Dialectical thinking and adult development.* Norwood, NJ: Ablex.

Basseches, M., & Mascolo, M. F. (2010). *Psychotherapy as a developmental process.* New York, NY: Routledge.

Commons, M. L., Goodheart, E. A., Pekker, A., Dawson, T. L., Draney, K., & Adams, K. M. (2007). Using Rasch scaled stage scores to validate orders of hierarchical complexity of balance beam task sequences. In E. V. Smith, Jr. & R. M. Smith (Eds.). *Rasch Measurement: Advanced and Specialized Applications* (pp. 121-147). Maple Grove, MN: JAM Press.

Commons, M. L., & Richards, F. A. (1995). Behavior analytic approach to dialectics of stage performance and stage change. *Behavioral Development, 5*(2), 7-9.

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

Commons, M. L., Rodriguez, J., Miller, P. M., Ross, S. N., LoCicero, A., Goodheart, E. & Danaher-Gilpin, D. (2007). *Applying the Model of Hierarchical Complexity.* Cambridge, MA: Dare Association, Inc. Unpublished manuscript available from commons@tiac.net.

Commons, M. L. & Ross, S. N. (Eds.). (2008). Special triple Issue on Postformal Thought and Hierarchical Complexity. *World Futures: Journal of General Evolution, 64*(5-7).

Commons, M. L., Ross, S. N., & Bresette, M. L. (2011). The connection between postformal thought, stage transition, persistence, ambition and major scientific innovations. In. C. Hoare, (Ed.), *Oxford handbook of reciprocal adult development and learning,2nd ed.* (pp. 287-301). Oxford University Press.

Commons, M. L., Trudeau, E. J., Stein, S. A., Richards, F. A., & Krause, S. R. (1998). The existence of developmental stages as shown by the hierarchical complexity of tasks. *Developmental Review, 8*(3), 237-278.

Fischer, K. W. & Yan, Z. (2002). Darwin's construction of the theory of evolution: Microdevelopment of explanations of variation and change in species. In N. Grannott & J. Parziale, (Eds.), *Microdevelopment: Transition processes in development and learning.* (pp. 294-318). Cambridge: Cambridge University Press.

Goldstein, J. A. (2002). The singular nature of emergent levels: Suggestions for a theory of emergence. *Nonlinear Dynamics in Psychology and Life Sciences, 6*(4), 293-309.

Goldstein, J. A. (2007). A new model for emergence and its leadership implications. In J. K. Hazy, J. A. Goldstein, & B. B. Lichtenstein (Eds.), *Complex systems leadership theory: New perspectives from complexity science on social and organizational effectiveness* (pp. 61-92). Mansfield, MA: ISCE Publishing.

Gelman, R., Romo, L., & Francis, W. S. (2002). Notebooks as windows on learning: The case of a science-into-ESL program. In N. Grannott & J. Parziale, (Eds.), *Microdevelopment: Transition processes in development and learning.* (pp. 269-293). Cambridge: Cambridge University Press.

Granott, N. & Parziale, J., (Eds.). (2002). *Microdevelopment: Transition processes in development and learning.* Cambridge: Cambridge University Press.

Kuhn, D. (2002). A multi-component system that constructs knowledge: Insights from microgenetic study. In N. Grannott & J. Parziale, (Eds.), *Microdevelopment: Transition processes in development and learning.* (pp. 109-130). Cambridge: Cambridge University Press.

Laske, O. (2009). *Measuring hidden dimensions of human systems: Foundations of requisite organization, Volume 2.* Medford, MA: IDM Press.

Parziale, J. (2002). Observing the dynamics of construction: Children building bridges and new ideas. In N. Grannott & J. Parziale, (Eds.), *Microdevelopment: Transition processes in development and learning.* (pp. 157-180). Cambridge: Cambridge University Press.

Richards, F. A., & Commons, M. L. (1990). Applying signal detection theory to measure subject sensitivity to metasystematic, systematic and lower developmental stages. In M. L. Commons, C. Armon, L. Kohlberg, F. A. Richards, T. A. Grotzer, & J. D. Sinnott (Eds.), *Adult Development*: Vol 2. *Models and methods in the study of adolescent and adult thought* (pp. 175−188). New York: Praeger.

Ross, S. N. (2007). *The fractal nature of stage change: A model and transition data.* (2007). Poster presented at the July 25 poster session of the Annual Meeting of Society for Mathematical Psychology, Costa Mesa, CA, July 25-28, 2007.

Ross, S. N. (2008). Fractal transition steps to fractal stages: The dynamics of evolution II. *World Futures: The Journal of General Evolution, 64*(5-7), 361-374.

Ross, S. N. & Commons, M. L. (2007). *The Fractal Nature of Stage Change: A Model and Transition Data.* Paper presented in Theoretical Foundations of Nonlinear Dynamics July 29, at 17th Annual International Conference of Society for Chaos Theory in Psychology & Life Sciences, Orange, CA, July 27-29, 2007.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press.

Sabelli, H. (1995). Non-linear dynamics as a dialectic logic. *Proceedings of the International Systems Society,* (pp. 101-112). Retrieved April 10, 2012 from http://www.creativebios.net/Nonlinea%20Dynamics%20as%20a%20process%20logic.html

Yan, Z. & Fischer, K. W. (2007). Pattern emergence and pattern transition in microdevelopmental variation: Evidence of complex dynamics of developmental processes. *Journal of Developmental Processes, 2*, 39-62.

BEHAVIORAL
DEVELOPMENT
BULLETIN

Toward defining order 16 and describing its performance for the model of hierarchical complexity

Sara Nora Ross[1], Michael Lamport Commons[2], Eva Yujia Li[3], Kristian Stålne[4], and Cory David Barker[1]

1 Antioch University Midwest
2 Harvard Medical School
3 Dare Institute
4 Department of Construction Sciences, Lund University

ABSTRACT

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.

KEYWORDS: cross-paradigmatic, meta-crossparadigmatic, model of hierarchical complexity, stage, performative-recursive

Fields of study have a social obligation to communicate about their findings as well as their state of evolution *as* areas of study. When areas of study or inventions within them are so new that the publication record does not yet report findings, to publish interim reports on the early discourse helps fulfill that obligation. This brief article falls in that genre of interim reporting. Our purpose is to offer a concise report on the first four years of the new theoretical discourse on the definition of order 16 of hierarchical complexity. In collaborating to report our progress on defining order 16 — including efforts to describe and measure tasks performed at that order—we move the theoretical discourse one step further in its evolution.

We begin by stipulating the meaning of *coordination,* a MHC term that is central for this discussion (other MHC terms are defined elsewhere in this issue). Tasks performed at an order of complexity *n* are actions that *coordinate* lower-order actions *n-1*. To *coordinate* means to *operate on*. These operations may take a range

of forms to: reflect on, compare, contrast, transform, define, and/or synthesize the *properties and behaviors of actions* (Commons, Ross, Miller, Richardson, Crone-Todd, & Miller, 2012; Ross, 2008). Note that "to understand" information is not one of the operations. This is because one can understand information at an order *n*, but could not have created the information nor coordinate it in a higher-order synthesis at *n+1*. In summary, Piaget's *operational* concept is central in this present discussion, as well as axiomatic in MHC theory: tasks of any order of complexity, *n,* operate on tasks performed at the *n-1* order of complexity *by coordinating* them.

» RECOGNIZING AN OCCURRENCE OF ORDER 16

The publicly-marked beginning of the discourse on order 16 was written in 2007 (published in late 2008), in the editors' introduction to the *World Futures* special issue on hierarchical complexity and postformal thought (Commons & Ross, 2008). That introduction traced the history of the MHC's development to that point, with the last entry in the history as follows.

Table 1. History of orders 14 and 15

Orders	Earlier work and sources	Ross (2008)
Order 14 paradigmatic	Descriptions: » Fit metasystems together to form new paradigms (2007 scoring manual) » Work with the relationship between very large and often disparate bodies of knowledge in order to reflect on, compare, contrast, transform, and synthesize multiple principles and metasystems. (2007 scoring manual & World Futures expansion) » Or show it is impossible to do so, if, in a domain, the highest stage task is showing that metasystems are incomplete and adding to them creates inconsistencies. No further stages in that domain on that sequence are then possible (Sonnert & Commons, 1994). » Definition: A paradigm is a systematized set of relations among metasystems that reflects a coherent set of assumptions (World Futures expansion)	Properties of structure and process (dynamics) that characterize disparate metasystems are seen to apply to or coordinate with one another. E.g., meta-system comparisons that describe paradigmatic relationships (slide 7, emphasis in original)
Order 15 cross-paradigmatic	Descriptions: » Fit paradigms together to form new fields (2007 manual) » Form new fields by crossing paradigms; » Integrate paradigms into a new field or profoundly transform an old one; » A field contains more than one paradigm and cannot be reduced to a single paradigm. » Definition: A cross-paradigm is a systematized set of relations among paradigms that reflects a coherent set of assumptions (World Futures expansion)	Properties of structure and process (dynamics) described by disparate paradigms are seen to apply across and operate on those paradigms (slide 7, emphasis in original)

Sara Ross is the one who pointed out that the model is fractal, since it shows by measuring any tasks that it is self-similar at all scales. She came up with the fractal characteristics of both the transition steps and within the smash sequence that is within the transitions.... Now, we have to come up with a name for the new order 16 in the Model. This is the stage-generator characteristic of the Model's axioms in action: to reflect on the tasks of a given order, one has to be performing at the next highest order. (Commons & Ross, 2008, p. 302)

Throughout this brief report we do not explicate examples or the coordinations that produced conclusions, and we have agreed to exclude mention of all but one or two examples. This is because such a project requires a separate paper, and as we report here, we have further to go before solid analysis can be offered. Thus, we include the foregoing excerpt only as the documented marker for the beginning of the order 16 work, without explaining how the work on the fractal dimensions of the model invokes order 16.

» REVISING ORDERS 14 AND 15

The next public marker was the March 2008 symposium of the Society for Research in Adult Development, where Ross (2008) reported her work to name and describe order 15, and to posit corrections to the MHC's descriptions of orders 14 and 15 (Table 1). Those corrections were necessary because the descriptions are the action building blocks: they represent the order *n* - 1 actions coordinated at the next higher order.

Ross proposed the revised descriptions to orders 14 and 15 to solve the earlier descriptive and definitional problems: "To date, the scoring manual's stage 14 and 15 descriptions (*a*) violate the content-free, scale-independence of hierarchical complexity its mathematical and fractal properties and (*b*) describe the task in terms of the content of social outcomes of performing the task (a field of study is a social outcome)" (2008, slide 6). This means the descriptions must be content-neutral and internally consistent with MHC as a general theory.

The distinction between task descriptions and orders' definitions seems crucial for our collective efforts. Presentations of the MHC have tended to rely on descriptions since the mathematical representation of hierarchical complexity includes no order-specific content. Definitions are qualitatively different from descriptions, of course, and descriptions need to be consistent with the related definitions. The work ahead involves agreed versions to describe the higher orders 14–16 and define their terms. As Barker (Personal Communication, 2012) stressed in one conversation, these will need to meet the test of representing all possible performances of each order *at all different scales* of task domains. This is challenging because of the vast number of task domains across orders. While some of us have looked to only the hard sciences for evidence of the highest orders of complexity, some of us argue such innovative performances are not confined to only that domain of human activity.

» NAMING AND DESCRIBING ORDER 16

Two names have been proposed for order 16. In the discussion below, Commons, Li, and Stålne use the term *Meta-cross-paradigmatic*. Ross proposed *Performative-Recursive* as a meaningful representation of the dynamics she had analyzed for several years, described as follows.

What the task performer is doing while embedded in the performance:

Observes and understands that by virtue of the cross-paradigms that account for their dynamics, disparate entities ranging from the universe, to paradigms, to species, to social metasystems, to individuals, for example, by their nature and/or with volition, perform recursive procession actions upon themselves, which transform them while and by performing each recursion; transformation may be "positive" or "negative." (Ross, 2008, slide 9, emphasis in original)

Subsequently, Ross and Barker became co-thinkers on scoring these dynamics and examples of them. They agree the description merits refinement and definitions need to be developed. To date, the performative-recursive name has held up its "goodness of fit" from their perspective. Commons, Li, and Stålne have not proposed a description for order 16, but instead, report on the process they went through in the attempt to do so.

» SCORING THE TASK OF DESCRIBING ORDER 16

By 2011, we all (Commons, Li, and Stålne) began to understand how string theory in physics might coordinate the two paradigms of quantum mechanics and the general theory of relativity. We reviewed the history of string theory and the task of coordinating these two lower order paradigms of quantum mechanics and relativity. Our description of such coordination from a developmental perspective of the field of physics was taken from string theory itself. We did not write the rules for such coordination explicitly because string theory states them and is too difficult to translate at this time into the MHC. So the performance is transitional to stage 15 at step 4 smash (see Commons & Richards, 2002). While we have had Stage 15 described since 1984, and Commons and Bresette (2000; 2006; Commons, Bresette, & Ross, 2008) have described many such historical examples of Stage 15, we failed to notice that it would take stage 16 (meta-cross-paradigmatic) to compare Stage 15 examples. (Also see the examples of stage 15, cross-paradigmatic in Stålne, Commons, and Li (in press). This last paper on new physics describes the integration of wave and gravity into string theory.) One has to consider whether or not a performance is stage 15 or not. That reflection requires one more stage of higher complexity as Dawn Schrader, (personal communication, 1985) pointed out in the early days of developing the MHC. At the stage 16, by defining and reflecting on the properties of stage 15 action, those actions point to the existence of new order/stage 16. To score material without matching to examples, one has to perform one stage higher than the material to be scored. So the performance to date is transitional to stage 16 at step 4 smash (Commons & Richards, 2002).

» DISCUSSION

While there is a reasonable measure of certainty that Stage 16 is attainable or may have already been attained, there are lingering issues that need to be addressed. One discussion point is in regards to Ross's proposed revised descriptions of the order 14 and 15. It could not have been known for certain that their previous definitions were in need of improvement until an adequate number of examples of such stages across multiple domains were compared. These new revisions improve the definitions to be more encompassing and robust. Similar to the definitions that came before them, these definitions need to be tested against the breadth of task actions across all domains. We might keep in the back of our minds the original definitions while the new definitions are tested individually by the adult development community, while looking towards a future time in which to reevaluate the revised definitions to see if another revision is needed.

Another point of discussion is, as mentioned above, there exist two proposed terms to be used for order/stage 16. For now, either of these terms may be used as placeholders until a final term is decided by participants of this discovery. But until a comparison of order 16 examples shows similar properties of what these magnitudes of task actions share in common, it may be too early to tell what term best fits its properties. Ross's description of order/stage 16 must be followed up by the aforementioned need for a written analysis. Commons and Stålne's careful documentation through the transition to 16 may be expected to result in a written analysis as well. Comparison of these analyses of transition into and attainment of Stage 16 may be joined by other analyses, all of which may lend to a future paper to continue the capture of the state of affairs of this endeavor. Authors of this paper, among others, are taking different approaches towards the transitions to, attainments of, and describing of order 16, which produces a much needed variety of approaches. We hope that individuals, in groups and in a larger collaborative effort, will yield the much needed data required to demonstrate Stage 16, and do so in an empirically testable way. Such an approach allows participants in this discussion and discovery to not only score their own task actions, but for participants' work to be scored by others to corroborate the scores to verify order 16 discovery and performances of tasks at stage 16.

Such checks and balances are especially important when coordinating more complex than usual hierarchical complexity and levels of abstraction. To bring this paper to a conclusion: this theoretical development terrain promises intense and promising work ahead on this breakthrough in the MHC, its contributions to behavioral development theory, and the measurement of the most complex human accomplishments recognized thus far. ∎

REFERENCES

Commons, M. L., & Bresette, L. M. (2000). Major creative innovators as viewed through the lens of the general model of hierarchical complexity and evolution. In M. E. Miller & S. Cook-Greuter (Eds.), *Creativity, spirituality, and transcendence: Paths to integrity and wisdom in the mature self* (pp. 167-187). Stamford, CT: Ablex Publishing Corporation.

Commons, M. L., & Bresette, L. M. (2006). Illuminating major creative innovators with the model of hierarchical complexity. In C. Hoare (Ed.). *Oxford handbook of development and learning*. (pp. 255-280). New York, NY: Oxford.

Commons, M. L., Bresette, L. M., & Ross, S. N. (2008). The connection between postformal thought and major scientific innovations. *World Futures: Journal of General Evolution, 64*, 503-512. doi: 10.1080/02604020802303838

Commons, M. L., & Ross, S. N. (2008). What postformal thought is, and why it matters. *World Futures: Journal of General Evolution. 65*(1-3), 321-329.

Commons, M. L., & Richards, F. A. (1984a). A general model of stage theory. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), *Beyond formal operations: Vol. 1. Late adolescent and adult cognitive development* (pp. 120-140). NY: Praeger.

Commons, M. L., & Richards, F. A. (1984b). Applying the general stage model. In M. L. Commons, F. A. Richards, & C. Armon (Eds.*), Beyond formal operations: Vol. 1. Late adolescent and adult cognitive development* (pp. 141-157). NY: Praeger.

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. & Ross, S. N. (2008). Editors' introduction to the special issue on postformal thought and hierarchical complexity. *World Futures: The Journal of General Evolution, 64*(5-7), 297-320.

Commons, M. L., Ross, S. R., Miller, P. M., Richardson, A. M., Crone-Todd, D. Miller, J. G. (2012). *Applying the model of hierarchical complexity: Scoring moral narratives and constructing moral development instruments* Unpublished manuscript available from commons@tiac.net.

Sonnert, G. & Commons, M. L. (1994). Society and the highest stages of moral development. *Politics and the Individual 4*(1), 31-55.

Stålne, K., Commons, M. L., & Li, E. Y. (In press). Hierarchical complexity in physics. *Journal of Adult Development.*

Ross, S. N. (2008). *Toward describing new order 15 in the model of hierarchical complexity and refining descriptions of paradigmatic order 13 and cross-paradigmatic order 14.* Paper presented at The Annual meeting of the Society for Research in Adult Development. March, 2008. Boston, MA.

The construction and validation of a developmental test for stage identification: Two exploratory studies

Hudson F. Golino[1], Cristiano Mauro Assis Gomes[1], Michael Lamport Commons[2], and Patrice Marie Miller[3]

1 Universidade Federal de Minas Gerais, Brazil
2 Harvard Medical School, USA
3 Salem State University, USA

ABSTRACT

The present work presents two exploratory studies about the construction and validation of the Inductive Reasoning Developmental Test (IRDT), a forty-eight items test based on the Model of Hierarchical Complexity. The first version of the test was administered to a convenience sample composed by 167 Brazilian people (50.3% men) aged between 6 to 58 years ($M = 18.90$, $SD = 9.70$). The Rasch Model was applied, and the result shows reliability of .97 for the full scale. The Infit mean was .87 ($SD = .28$; Max = 1.69; Min = .39), and the person reliability was .95. The one sample t-tests showed significant spacing of Rasch scores between items of adjacent orders of hierarchical complexity, with large effect size. The second study was conducted in order to overcome some of the test's limitations found in the first study. The revised IRDT were administered to a convenience sample composed of 188 Brazilian people (57.7% women) aged between 6 and 65 years ($M = 21.45$, $SD = 14.31$). The reliability for the full scale was .99, and its Infit mean was .94 ($SD = .22$; Max = 1.46; Min = .56). The person reliability was .95. The one sample t-tests showed significant spacing of Rasch scores between items of adjacent orders of hierarchical complexity, with large effect size. The paper finishes with a discussion about the necessity and importance to focus on the vertical complexity of the items in any test designed to identify developmental stages.

KEYWORDS: stages, assessment, validation, development, model of hierarchical complexity, inductive reasoning

Piaget is considered one of the most important researchers of the 20th century (Flavell, 1963), with his studies creating a very influential framework within developmental psychology, that of Genetic Epistemology. In spite of its importance, the influence of this theory on developmental research began to decline in the 1980's, due to a large body of evidence that apparently contradicted the theory's notion of developmental stages (Marshall, 2009; Miller, 2002). One might say that this theory was "put in check" by the maneuvers of others. When Piaget's theory, specifically his stage concept, was put in check, all Piagetian and Neo-Piagetian developmentalists were, in some manner, placed in the same condition. As in chess, getting out of the check is of great importance, and requires the development and implementation of

sturdy strategies. In developmental psychology, getting out of check can be reached through the implementation of "strategic moves", as in the construction of better metrics (Fischer & Rose, 1999; Rose & Fischer, 1998; Van Geert & Steenbeek, 2005), with reliable, valid and accurate measures (Fischer & Dawson, 2002), and the adoption of quality control standards (Stein & Heikkinen, 2009).

The current paper presents one of these moves which, together with other works (Commons, Trudeau, Stein, Richards, & Krause, 1998; Commons et al., 2008; Dawson, 2003, 2006; Dawson & Wilson, 2004; Dawson, Goodheart, Wilson, & Commons, 2010; Dawson-Tunik, Commons, Wilson, & Fischer, 2005; Demetriou & Kyriakides, 2006; Fischer, 2008; Fischer & Bidell, 1998, 2006; Rijmen, De Boeck, & Van der Mass, 2005; Van der Maas & Molenaar, 1992), aims to collaborate in getting out of the check. Two exploratory studies about the construction, challenges and initial results of the Inductive Reasoning Developmental Test (IRDT) - *Teste de Desenvolvimento do Raciocínio Indutivo* (Gomes & Golino, 2009) will be presented. The IRDT intends to measure the developmental stages of inductive reasoning through reliable, valid and accurate measures, falling in the category of so-called "quality control standards".

Criticisms of stages, or killing Piagetian stage theory:

Starting in the 1980's, increasing numbers of researchers began to criticize Piagetian stage theory (Miller, 2002; Morra, Gobbo, Marini, & Sheese, 2008). The main criticisms were directed at the idea that stages are structures of the whole, developing in a synchronous way, emerging at specific ages, and reaching a single *telos*, represented by formal operations (Fischer & Bidell, 2006).

One set of criticisms that emerged empirically supported the idea that variability is the norm, rather than the exception in human development (Bidell & Fischer, 1992, 2006; Fischer & Rose, 1999; Flavell, 1963; Miller, 2002; Siegler, 1981). Such evidence points to asynchrony, heterogeneity and high variability in performance (Demetriou, Efklides, Papadaki, Papantoniou, & Economou, 1993; Fischer & Bidell, 2006). Some major studies indicate *decaláge* in the ability of seriation (Chapman & Lindenberger, 1988; Halford, 1989; Jamison, 1977), conservation (Kreitler & Kreitler, 1989; Nummedal, 1971; Murray, 1969; Murray & Holm, 1982), formal operations (Bart, 1971; Lautrey, de Ribaupierre & Rieben, 1985; Martorano, 1977; Webb, 1974), combinatorial analysis (Roberge, 1976; Scardamalia, 1977), object permanence (Baillargeon, 1987; Chazan, 1972; Jackson, Campos & Fischer, 1978), among others.

In addition to studies showing massive *decaláges*, age issues and synchronism problems in Piagetian theory of cognitive development, other revisions of the theory were made. Commons and Richards (1984a), Commons, Richards and Kuhn (1982), Fischer (1980, 1987), Fischer, Hand and Russell (1984), and others, argued that the stage of formal operations is not the last possible level in human cognitive development, and show evidence for post-formal levels.

The other set of criticism emerged from philosophical/epistemological positions. Broughton (1984), for example, argued that formal operations are a wholly inadequate model of thought in adolescence and adulthood, and as a result suggests the entire theory should be reconsidered.

The criticism, sometimes based on empirical aspects, sometimes based on philosophical and epistemological positions, was striking, and came from many different lines. Flavell already in his early work entitled *The Developmental Psychology of Jean Piaget* (1963), pointed to ambiguities in the concept of stage, argued about the challenges of the clinical method, on the impossibility of stating that a child "has" a particular concept and raised the question of language as an intervening variable (Siegler & Crowley, 1991). Despite recognizing the historical importance of Piaget's work, in particular the stage theory, Flavell comes to argue, in another, later work, that the Piagetian stage theory "explains nothing" (Flavell, 1985; Lourenco, 1998). Lourenço (1998) proposed that many cognitivists (e.g., Bjorklund, 1997; Brainerd, 1997; Cohen, 1983) already considered Piaget's theory to be dead, and some of them suggested that there was no real purpose in continuing to test a theory that was already known to be inadequate (Halford, 1989; Lourenco, 1998).

In short, until the mid 80's the classic structuralism of Piaget's theory had significantly influenced developmental psychology research worldwide (Marshall, 2009). In spite of being one of the most important players of the "Developmental Chess," the grandmaster was double *checked*. His influence, including the concept of stages, began to decline, due mainly to (1) the growing body of evidence that helped convince some researchers that stage theory was inappropriate to describe cognitive development (Morra, et al., 2008), and to (2) criticisms that addressed philosophical issues and suggested an epistemological reconfiguration (Marshall, 2009).

Neo-Piagetians and Post-Piagetians

A group of Neo-piagetian researchers has sought to overcome the problems and limitations pointed to in the Piagetian concept of stage, including his methodology for assessing them, proposing instead modern theoretical and methodological approaches that have been providing new evidences for discontinuity. Included in these new approaches are two important and related models of development: Fischer's Dynamic Skill Theory (DST; Fischer, 1980; 2008) and Commons' Model of Hierarchical Complexity (MHC; Commons, 2008; Commons & Richards, 1984; Commons & Pekker, 2008). Fischer (1980) proposed a set of analytical tools that make possible the detailed description of developmental pathways, as well as the construction of domain-free hierarchical taxonomies to classify performance. His DST (Fischer, 1980; 2008; Fischer & Bidell, 1998, 2006; Fischer & Rose, 1994, 1999; Fischer & Yan, 2002a, 2002b) conceives of development as a phenomenon composed of both continuous and discontinuous patterns of changes. The former (continuous change) relates to the sequence of steps followed in the construction of skills (microdevelopment) and the latter (discontinuous change) relates to abrupt, stage-like changes that marks the emergence of radically new kinds of control units of behavior and cognition (Fischer, 1980; Fischer & Rose, 1994; Fischer & Bidell, 1998, 2006; Fischer & Yan, 2002a). Evidence for both kinds of developmental patterns have been provided by Fischer and colleagues (Fischer, Kenny, & Pipp, 1990; Fischer & Silvern, 1985; Fischer & Yan, 2002a, 2002b; Schwartz & Fischer, 2005; Yan & Fischer, 2007). Instead of conceptualizing the discontinuous facet of human development as a unidirectional ladder the DST sees it as a *constructive web* that encompasses someone's activity and the supportive context in which this activity is performed (Bidell & Fischer, 1992; Fischer & Bidell, 2006). So, a person may have a certain level of performance, let us say *x*, in the domain of Algebra, and an *x*-1 level of performance in the domain of Combinatorial Analysis, for example. Furthermore, this same person may present higher or lower levels of performance in the previously cited domains due to social support (scaffolding), emotional reactions, and so on (Fischer & Bidell, 2006). The constructive web notion is different from the Piagetian concept of stages as developmental ladder, in which *decaláge* is the exception.

Despite the importance and contribution of the DST to the Developmental Sciences field (Miller, 2002; Morra et. al, 2008), it was Commons and his colleagues that have proposed the groundwork for the mathematical formalization of discontinuity, through the Model of Hierarchical Complexity (MHC). The MHC is a general measurement theory, and as such is part of the normal Mathematical Theory of Measurement (Krantz, Luce, Suppes, & Tversky, 1971; Luce, & Tukey, 1964) applied to the phenomenon of difficulty. The MHC introduces the concept of the Order of Hierarchical Complexity (OHC) that conceptualizes information in terms of "the power required to complete a task or solve a

problem" (Commons, Trudeau, Stein, Richards, & Krause, 1998). Commons and Pekker (2008) demonstrated, in axiomatic terms, that task difficulty or complexity, beyond other sources, increases in two ways: horizontally and vertically. The first refers to the accumulation of informational bits necessary to successfully complete a task (Commons, 2008), e.g. $5 + 6 + 7$ is less complex than $5 + 6 + 7 + 8$, because the first differs from the second in the number of times addition was executed, and does not differ in the organization of the addition itself; that is, both have the same *hierarchical (or vertical) complexity.* So, horizontal or traditional complexity is just the adding of informational bits. Vertical complexity, or *hierarchical complexity*, refers to the organization of information in the form of action in two or more subtasks, in a coordinated way. The distributive property is a good example of vertical complexity. Let's take the following example: $a \times (b + c) = (a \times b) + (a \times c)$. In order to correctly perform the task, one should multiply the element *a* by *b* and by *c*, separately, and then sum the results, or sum *b* with *c*, and then

Table 1. Some instruments based on the model of hierarchical complexity and/or dynamic skill theory

Problem-solving

Algebra (Richardson & Commons, 2008)

Balance beam (Dawson, Goodheart, Draney, Wilson, & Commons, 2010)

Infinity (mathematics) (Richardson & Commons, 2008)

The laundry problems (Goodheart & Dawson, 1996; Goodheart, Dawson, Draney, & Commons, 1997)

The combustion problem (Bernholt, Parchmann, & Commons, 2008).

Vignettes

Social perspective-taking (Commons & Rodriguez, 1990; 1993)

Informed consent (Commons & Rodriguez, 1990, 1993; Commons, Rodriguez, Adams, Goodheart, Gutheil, & Cyr, 2006)

Attachment and loss (Miller & Lee, 2000)

Workplace organization (Bowman, 1996a; 1996b)

Workplace culture (Commons, Krause, Fayer, & Meaney, 1993)

Political development (Sonnert & Commons, 1994)

Relationships (Armon, 1984a)

Views of the "good life" (Danaher, 1993; Dawson, 2000; Lam, 1994)

Epistemology (Kitchener & King, 1990; Kitchener & Fischer, 1990)

Moral judgment (Armon & Dawson, 1997; Dawson, 2000)

The Helper-Person Problem, The Incest Dilemma Against, The Pro-Death Penalty Dilemma, The Anti-Death Penalty Dilemma, The Politician-Voter Problem, The Christ Stoning Case Without Sin (Miller, Bett, Ost, Commons, Day, Robinett, Ross, Marchand, & Lins, 2008)

Other

Four story problem (Commons, Richards & Kuhn, 1982; Kallio & Helkama, 1991)

Counselor stages (Lovell, 2002)

Loevinger's sentence completion task (Cook-Greuter, 1990)

Report patient's prior crimes (Commons, Lee, Gutheil, Goldman, Rubin, Appelbaum, 1995)

Causing religious beliefs / causing atheism (Miller, Harrigan, Commons, & Commons-Miller, 2008)

The student-bully problem (Joaquim, 2011)

multiply by *a*. If someone change the order of execution of the actions, e.g. $(a \times b) + c$, the result won't be right. So, it requires the two actions of addition and multiplication to be performed in a certain order, thus, coordinated.

Briefly summarizing, the MHC postulates that actions at a higher order of hierarchical complexity: 1) are defined in terms of two, or more, lower-order actions; 2) organize and transform those actions, not just combine them in a chain; and 3) produce organizations of lower-order actions that are new and not arbitrary. The first two are also Piagetian postulates, but the third is not. The order of hierarchical (or vertical) complexity refers to the number of recursions that the coordinating actions must perform on a set of primary elements (Commons, 2008). Because hierarchical complexity is a property of tasks, performance is separated from tasks. Stage is defined as the most hierarchically complex task solved. Each task that occurs in a separate domain is considered separately. There is no structure of the whole, so in the DST, *decaláge* is the normal modal state of affairs.

The development of metrics in developmental psychology has been one of the challenges and needs of the area (Van Geert & Steenbeek, 2005; Fischer & Rose, 1999), and it is considered crucial in guiding research and professional practice (Stein & Heikkinen, 2009). The Hierarchical Complexity Score System – HCSS (Commons, LoCicero, Ross & Miller, 2010; Dawson, Commons, Wilson,

& Fischer, 2005) and the Lectical Assessment System – LAS (Dawson-Tunik, 2004) represent general, reliable, valid, domain-free scales or metrics (Dawson, 2004). These metrics were studied by Dawson (2000, 2001, 2002, 2003, 2004) who compared them with domain-specific scales, such as the *Good Life Scoring System* (Armon, 1984), the *Standard Issue Scoring System* (Colby & Kohlberg, 1987a, 1987b) and the *Perry Scoring System* (Perry, 1970). Dawson (2003) points out that, in spite of measuring the same latent variable, the domain-free scales present better internal consistency, allow meaningful comparisons across domains and contexts, and enable the examination of the relationship between developmental stages and conceptual content. Moreover, the HCSS and the LAS are considered two of few calibrated developmental metrics in use, being studied in terms of their construct and congruent validity, internal consistency and inter-rater reliability, providing evidence of fine grained interval scales (Stein & Heikkinen, 2009).

Despite the importance in guiding developmental and psycho-educational research and practice, the domain-specific scales demand various trained scoring analysts, with high agreement between them, require a considerable time for large scale evaluation and are vulnerable to subjective bias. So, the construction of objective large-scale tests can help the field to move beyond these challenges, bringing speed and lower-cost procedures for evaluating discontinuities.
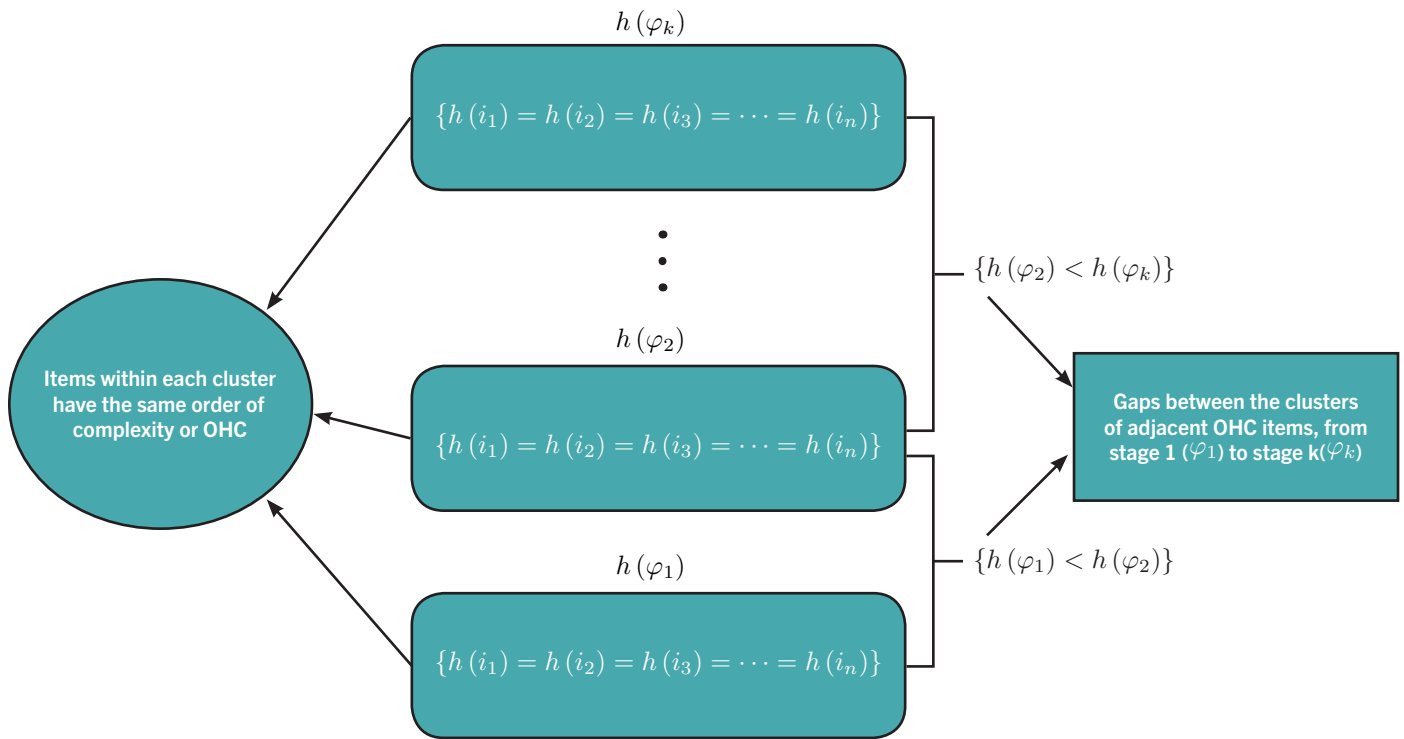
Figure 1. Expected item structure of instruments constructed focusing on the vertical complexity within a specific domain (unidimensional)

The MHC can be used not only to construct analytic scales, but also for the construction and design of tests, tasks and vignettes. A number of tasks have been created in many domains, based on the MHC or DST (as seen in Table 1).

Constructing calibrated tests for developmental stage identification requires a specific design as defined by Commons and colleagues (Commons & Pekker, 2008; Commons, Gane-McCalla, Barker & Li, this issue). This design involves: 1) grouping items with same hierarchical complexity $[h(i_1) = h(i_2) = h(i_3) = … h(i_n)]$ within stages; and 2) using items with increasing hierarchical complexity $[h(\text{Stage 1}) < h(\text{Stage 2}) < h(\text{Stage 3}) < … h(\text{Stage } k)]$ between stages. The first deals with item or task equivalence, important in order to avoid the elaboration of an anomalous scale that confuses its analysis (Fischer & Rose, 1999). The second makes possible the identification of discontinuous, stage-like development, with gaps between different orders. There is an expected item structure of any instrument construct based on the MHC. That structure focuses on both strategies in order to identify developmental stages should be as close as possible to the diagram below (Fig. 1). Each blue box in the Figure 1 represents a cluster of items of the same unidimensional domain. Within a single box, the items have the same Order of Hierarchical Complexity (h) in that domain. The OHC of the items increases from stage 1 (φ_1) to stage k (φ_k), so that $h(\varphi_1) < h(\varphi_2) < \cdots < h(\varphi_k)$ (Consequences 2, 3 and 4 of the formal MHC; see Commons & Pekker, 2008; Commons, Gane-McCalla, Barker & Li, this issue). Furthermore, the figure shows the expected gaps between the clusters of adjacent OCH items (see Figure 1).

Beyond the strategies of grouping items with same OHC and using items with increasing OHC, in order to identify developmental stages, a good measure or ruler needs to address a single trait or dimension, be constructed based upon an explicit theory or model of development (Stein, Dawson & Fischer, in press), be submitted to empirical investigation, aiming to test the expected equivalence and order of items, and determine other scale properties (Fischer & Dawson, 2002; Fischer & Rose, 1999). Commons and colleagues (Commons et al., 2008; Dawson, Goodheart, Draney, Wilson, & Commons, 2010) evaluated the expected equivalence and order of items from the developmental test design through the Rasch family of models (Andrich, 1988; Rasch, 1960). The dichotomous Rasch Model (Rasch, 1960/1980), also called Simple Logistic Model (SLM) for dichotomous responses (Andrich, 1988), establishes that the right/wrong scored response x_{vi}, that emerges from the encounter between the person v and the item i, depending upon the performance β of that person and on the difficulty δ of the item. Its relation can be expressed as the following probabilistic function:

$$P\{X_{vi} = x\} = \frac{e^{x(\beta_v - \delta_i)}}{1 + e^{(\beta_v - \delta_i)}} \qquad (1)$$

The Rasch model deals with the relationship between the person ability and item difficulty in a probabilistic way. Both parameters are allocated on a single abstract continuum that goes from "low" to "high" ("more" or "less", etc), concerning just one attribute of the object (or attitude, or behavior) measured, thus *unidimensional*. In the Classical Test Theory (CTT) the corresponding "parameter" for the Rasch's person performance (β_v) is the estimated *true score* $(\hat{T}_v)$, or the score reported on test-score scale (normally distributed) (Hambleton & Jones, 1993). It can indicate the "position" of the person on the construct measured, but unlike the SLM, needs a representative sample for unbiased item estimates, a norm group

for comparison between individuals, giving meaning to the scores, and a normally distributed score for achieving interval scales properties (Embreston & Reise, 2000).

Some authors argue that the dichotomous Rasch model is the simplest Item Response Theory model (one-paramenter model) (Bock & Jones, 1968; Hambleton, 2000). However, Andrich (2004) argues that differently from the traditional IRT paradigm, in which one chooses the model to be used (one, two or three parameters) according to which better accounts for the data, in the Rasch Paradigm "the SLM is used because it arises from a mathematical formalization of invariance which also turns out to be an operational criterion for fundamental measurement" (p.15). So, instead of data modeling, the Rasch's paradigm focuses on the verification of data fit to a fundamental measurement criterion, compatible with those found in the physical sciences (Andrich, 2004. p.15).

From among the benefits of using the Rasch family of models for measurement, some should be highlighted. In sum, it allows the construction of objective and additive scales, with equal-interval properties (Bond & Fox, 2001; Embreston & Reise, 2000), it produces linear measures, gives estimates of precision, allows the detection of lack of fit or misfit and enables the parameters' separation of the object being measured and of the measurement instrument (Panayides, Robinson & Tymms, 2010). It also makes possible the reduction of all of a test's items into a common developmental scale (Demetriou & Kyriakides, 2006), collapsing in the same latent dimension person's abilities and item's difficulty (Bond & Fox, 2001; Embreston & Reise, 2000; Glas, 2007), and enables the verification of hierarchical sequences of both item and person, being especially relevant to developmental stage identification (Dawson, Xie & Wilson, 2003).

Through the assumptions and procedures introduced by Commons and colleagues (Commons and Pekker, 2008; Commons et al., 2008; Dawson-Tunik et al., 2010) it has become possible to design and construct valid and reliable developmental metrics, tests and tasks, bringing new empirical evidence that helps reveal stage-like discontinuity. Following this tradition, two exploratory studies about the construction, challenges and initial results from the construction of an objective, large-scale instrument, named the Inductive Reasoning Developmental Test (IRDT), developed by Gomes and Golino (2009). These studies will be presented in some detail with the aim of unpacking the challenges involved in the construction of a developmental test, and will present a methodology for developmental stage identification. This methodology is put forward as one of the moves that can help uncheck the idea of stages within the virtual game of "Developmental Chess", together with other moves published elsewhere (Demetriou & Kyriakides, 2006; Rijmen, De Boeck, & Van der Mass, 2005).

Study I: Uncovering discontinuities, and finding alternative sources of difficulty beyond vertical complexity

The purpose of Study 1 was to construct the initial version of the instrument, and in so doing, assess the scale structure of the items, verifying if they presented previously predicted orders and gaps, and to investigate the initial estimates of reliability and unidimensionality, among other scale properties, using Rasch analysis.

1	A	A	A	A	E

Figure 2. Example: Item 1, Stage Pre-op.

The IRDT (Gomes & Golino, 2009) is a pencil-and-paper instrument designed to assess developmentally sequenced and hierarchically organized inductive reasoning. It is an extension, in terms of complexity, from the *Indução* test, which compose the fluid intelligence test kit (Gomes & Borges, 2009) of the Higher-Order Cognitive Factors Kit (Gomes, 2010). The domain of inductive reasoning was used because it is one of the best indicators of fluid intelligence (Carroll, 1993). The construction of the IRDT, from the original *Indução* items, is due to a larger challenge that concerns the construction of an intelligence battery to identify developmental stages.

The sequence of IRDT was constructed based on the MHC and on Fischer's Dynamic Skill Theory. It was designed to identify six developmental stages (or levels), that will be named based in both theories, respectively: Pre-operational or Single Representations (Pre-op/SR); Primary or Representational Mappings (Prim/RM); Concrete or Representational Systems (Conc/RS); Abstract or Single Abstractions (Abst/SA); Formal or Abstract Mappings (Form/AM); and Systematic or Abstract Systems (Syst/AS). Each stage is composed of eight items with the same order of hierarchical complexity (OHC), for a total of forty-eight items. Each item is composed of four letters, or sequence of letters, with a specific rule (correct items), plus one letter or sequence with a different rule (exception). The task is to discover which letter or sequence is the exception. From stage to stage, there is a difference of +1 in the Order of Hierarchical Complexity (OHC). The instructions for performing the test is as follow: "You'll be presented several reasoning tasks (items). In each task (item) you have five letters or sequence of letters. Among the five letters or sequence of letters, four of them have a specific rule, and one has a rule that is different from the others. Your challenge is to identify (marking with an X) the letter or the sequence of letters that has a different rule, compared to the other four. Each task (item) is displayed in a specific row, beginning with a number, from 1 to 48. You have no time limit. Solve as many tasks (items) as you can."

Pre-operational or Single Representations (Pre-op/SR): Each item is composed of specific letters. The rule is "equal letter", and the exception is a different one (see Figure 2).

Primary or Representational Mappings (Prim/RM): Eight items were created for this stage. Four of them have a specific rule: there is no jump in the letters' sequence. In the example below, the first option is composed of WX. There is no other letter between them, so they form a non-jump sequence (Rule 1). The exception, however, is a conjoint of two letters that jumps one letter of the alphabetic sequence (e.g. QS; see Figure 3).

9	WX	MN	ST	QS	YZ

Figure 3. Example: Item Prim/MR1 — Rule 1

13	XZ	DF	MO	RU	HJ

Figure 4. Example: Item 13, Primary/MR — Rule 2

17	IJKM	NOPR	EFHI	QRSU	JKLN

Figure 5. Example: Item 17, Concrete/RS — Rule 3

The other four items of the Primary Stage follows the same structure, but have different rules. The majority of the options jump one letter of the alphabetic sequence (Rule 2). So, in the example below, the option DF jumps the letter E. The exception is a conjoint of two letters that jumps two letters of the alphabetic sequence (e.g. RU; see Figure 4).

Concrete or Representational Systems (Conc/RS): All items are composed of four sets of four letters with one of the three following rules. In Rule 3 there is a jump of one letter only between the last two letters. For one example, see the item below. Between I and J, and between J and K, there is no other letter. However, there's a jump between K and M. The exception, in this item (17), is represented by the sequence EFHI, where the jump is located between the two letters in the middle (FH; see Figure 5).

In Rule 4, the jump occurs between the first pair of letters, and the exception is the option where the jump occurs between the two middle letters. The example below shows item 20. Note that the option NPQR presents a jump between N and P, like three other options. However, the first option (KLNO) presents a jump between the two middle letters, i.e. L and N (see Figure 6).

Finally, in rule 5 the jump occurs twice, between the two first pairs of letters. In the exception, the jumps occur between the first pair and between the last pair of letters. See the example below. In item 22, in the first option (RTVW) there is a jump between R and T, and between T and V, as in three other options. However, in the option BDEG, the jumps occur between B and D, and E and G (see Figure 7).

So, the first two items (Prim/RS1 and Prim/RS2) use rule 3, the items Prim/RS3 and Prim/RS4 use rule 4, and the other four items use rule 5.

Abstract or Single Abstractions (SA): Different from all other stages, here a table is introduced with codes referring to a coordination of two sets of four letters, in which the rules and exceptions presented at the Concrete/SR's items are also coordinated, forming new rules and exceptions. This coordination is shown by the *plus* sign between the letter sequences (see Figure 8).

The table has eight code rows, each beginning with an alphabetic letter followed by a Greek letter. So, the first code row has letter A followed by different Greek letters, while the second code row has letter B followed by the same Greek letters, and so on (see Figure 9).

The item to be answered is composed only by the table codes, in sequence. For example see Figure 10.

Formal or Abstract Mappings (Form/AM): All items are composed of a coordination of two codes, based on those presented at the Abstract Stage's table (see Figure 11).

Systematic or Abstract Systems (AS): All items are composed by a set of four codes, based on the previous presented at Abstract Stage's table (see Figure 12).

All items of the same stage were presented together at a specific page, so different stages were in different pages. The alphabetic sequence (all letters from A to Z) were printed above the items in each page, for consultancy. The order of hierarchical complexity is represented in the figure 13 below. It is important to note that the revision of the MHC stages was not incorporated in this study. The OHC numbers presented in this study are the older version of MHC stages. In the newer version, the numbers increase by one.

The Systematic items (OHC 11) coordinate two formal (OHC 10) components. By its turn, the formal items coordinate two abstract (OHC 9) components. The abstract items coordinate two concrete (OHC 8) components. The concrete items coordinate two primary (OHC 7) components. Finally, the primary items coordinate two pre-operational (OHC 6) components (see Figure 13).

» METHOD

Participants

In Study 1, the IRDT was administered to a convenience sample composed by 167 Brazilians (50.3% men, 49.7% women) aged between 6 to 58 years ($M = 18.90$, $SD = 9.70$). The sample was intentionally broad, and had a distribution of 15.6% from 6 to 12 years, 27.5% from 13 to 15 years, 35.9% from 16 to 20 years, and 21% beyond 20 years. All the participants were from the city of Belo Horizonte, state of Minas Gerais, Brazil.

Procedure

The data were collect by the first author and by thirty Psychology undergraduate students, enrolled in a first semester Cognitive Development class, trained in how to administer the instrument properly. The author first administered the instrument to the undergraduate students (whose data are being used in this analysis), and to 47 first year high school students from a public school. Each undergraduate student was assigned to administer the IRDT to three different people from 6 to 60 years of age. Participation was voluntary, with people agreeing to be part of the study after its purpose was explained. They were informed that their answers would be kept confidential, and that all procedures guaranteeing the privacy of their results would be adopted. They then signed an inform consent form, as required by the guidelines of the Ethical Committee of the Universidade Federal de Minas Gerais, Brazil.

Data analysis

In the first part of the data analysis the dichotomous Rasch Model is used, making it possible to reduce the items from the IRDT into a developmental scale (Demetriou & Kyriakides, 2006), collapsing at the same level person's abilities and item's difficulty (Bond & Fox, 2001; Embreston & Reise, 2000; Glas, 2007). It also enables

20	KLNO	NPQR	QSTU	DFGH	HJKL

Figure 6. Example: Item 20, Concrete/RS — Rule 4

22	RTVW	ACEF	CEGH	FHJK	BDEG

Figure 7. Example: Item 22, Concrete/RS — Rule 5

Table 2. One-sample tests of mean item difficulties for different OHC's

	Test value = 0
			Sig.			\multicolumn{2}{c}{95% CI}	\multicolumn{2}{c}{Effect size}
Stages	t	DF	(2-tailed)	M	SD	Lower	Upper	(d')	(r)
Pre-op/SR and Primary/RM	13.58	7	0.00	3.82	0.80	3.15	4.48	10.26	.98
Primary/RM and Concrete/RS	3.29	7	0.01	2.18	1.87	0.61	3.74	2.49	.77
Concrete/RS and Abstract/SA	7.99	7	0.00	1.69	0.60	1.19	2.18	6.03	.94
Abstract/AS and Formal/AM	36.01	7	0.00	2.89	0.23	2.70	3.08	27.28	.99
Formal/AM and Systematic/AS.	9.49	7	0.00	2.28	0.68	1.71	2.85	7.17	.96

the verification of hierarchical sequences of both item and person, being especially relevant to developmental stage identification (Dawson, Xie & Wilson, 2003).

To verify the adjustment of the data to the model, the Infit (information-weighted fit) mean-square statistic is used. It represents "the amount of distortion of the measurement system" (Linacre, 2002. p.1). Values between 0.5 and 1.5 logits are considered productive for measurement, and <0.5 and between 1.5 and 2.0 are not productive for measurement, but do not degrade it (Wright & Linacre, 1994). The unidimensionality of the instrument can be checked by a number of procedures, each one complementing the other (see Tennant & Pallant, 2006). Here, unidimensionality will be addressed using only the model fit statistics – i.e. if the data fit the model, one of the consequences is the linearity of the measure, its unidimensionality, and so on – and the principal contrast, which can be verified through the percentage of variance explained by measures, and by the percentage of unexplained variance in the first contrast. The former should be closer to or greater than 60% (Peeters & Stone, 2009), while the latter should be closer to or less than 10%.

In the second part of the analysis, the spacing of Rasch scores between items of adjacent orders of hierarchical complexity is described. The Rasch scores represent the difficulty of an item (δ), which is its location at the latent variable continuum. It would have been good to compare the Rasch Scores for every item from adjacent orders of hierarchical complexity, but because there were so many items, this would have produced too many comparisons. To reduce the number of comparison pairs, each item's Rasch score was subtracted from the mean Rasch score of the items from the next higher order of complexity. This calculation is represented by the Formula 2:

$$\overline{X}_{k+1} - \delta_{i_k} = Adj\delta_{i_k} \qquad (2)$$

where $\overline{X}_{k+1}$ is the mean of the next higher order of complexity (or Stage $k+1$), and δ_{i_k} is the difficulty of item i from order k (or Stage k) , producing the adjusted difficulty of item i. To verify if the differences between difficulties of items from order k and the mean difficulty of the order $k+1$ are statistically significant, the

One-Sample t-test is used, with a 95% confidence interval. The effect size (ES) is calculated using the Cohen's d and effect size correlation r (Rosnow & Rosenthal, 1996).

» RESULTS

The Rasch dichotomous model (Andrich, 1988; Rasch, 1960) was calculated using the software Winsteps (Linacre, 1999, 2011). Out of the 48 items, 5 were responded correctly by all participants (Pre-op/SR1, Pre-op/SR3, Pre-op/SR4, Pre-op/SR5 and Pre-op/SR8). The reliability for the forty-three non-extreme items was .99, and for the full scale (48 items) the reliability was .97. The Infit mean was .87 (SD = .28; Max = 1.69; Min = .39), falling within the acceptable fit range. The person reliability was .95, which is estimated to indicate the degree to which a person's response pattern conforms to the difficulty structure of the measure (Hibbard, Collins, Mahoney & Baker, 2009). The principal contrast showed that the raw variance explained by measures (modeled) is 70.6%, and that the unexplained variance in the first contrast (modeled) is 10.4%, suggesting that the instrument can be thought of as unidimensional.

The variable map (Figure 14) illustrates the scale for the 48 items of the IRDT with item difficulties (on the right) and person measures (on the left) calibrated on the same scale. It is visually possible to identify clear item clusters in the Systematic/Abstract Systems' stage (Syst/AS1, Syst/AS2, Syst/AS3, …, Syst/AS8) and in the Formal/Abstract Mappings's stage (Form/AM1, Form/AM2, Form/AM3, …, Form/AM8), with a gap between them. The Abstract/ Single Abstraction's items presented a cluster (they are all together without any other stage's items), but did not present a gap in relation to the Concrete/Representational System's items. Some Primary/Representational Mapping's items (Prim/RM5, Prim/RM6, Prim/RM7, Prim/RM8), had difficulties very close to the Concrete/RS's items, making one big item set. The other Primary/RM's items (i.e. Prim/RM1, Prim/RM2, Prim/RM3 and Prim/RM4) were less difficult than other items of the same stage. Moreover, they presented a gap in relation to the item's set composed by the other Primary items and by the Concrete ones. Finally, the relative position of person (left) and item (right), shows the IRDT as an easy test for 23 participants (Mean ability = 7.66, SD = 0.81). The whole sample mean ability was 1.15 with standard deviation of 3.40 logits (see Figure 14).

Aπ	Aδ	Aη	Aμ	Aλ
FGIK+OQST	OPRT+DFHI	IJLN+PRSU	EFHJ+TVXY	STVX+NPRS

Figure 8. Example: Table Row 1, Abstract/SA

Bπ	Bδ	Bη	Bμ	Bλ
QRTV+MOQR	UVXZ+FHJK	HIKM+SUWX	CDFH+NORS	GHJL+PRTU

Figure 9. Example: Table Row 2, Abstract/SA

25	Aπ	Aδ	Aη	Aμ	Aλ

Figure 10. Example: Item 25, Abstract/SA

33	AδFπ	CηEπ	BδEλ	CδFη	AλBπ

Figure 11. Example: Item 33, Formal/AM

The one-sample *t*-test, with 95% confidence interval, shows that the comparisons of difficulty between Pre-operational and Primary, Primary and Concrete, Concrete and Abstract, Abstract and Formal, and between Formal and Systematic were significant. Moreover, effect size *d'* and *r* were large (see Table 2).

» DISCUSSION

The current study aimed to assess the scale structure of the items, verifying whether they represented previously predicted orders and gaps (see Fig.1), and to investigate the initial estimates of reliability and unidimensionality, among other scales properties, using Rasch analysis. The result suggests the unidimensionality of the items, to some extent, since the percentage of raw variance explained by the measures (modeled) is moderately high (70.6%), and the principal components analysis of the residuals gave an unexplained variance of 10.4% for the first contrast. The items' adjustment to the model was verified through the Infit index, which was found to have a mean of .87 and a standard deviation of .28. The minimal Infit value was .39 (Item System/AS4) and the maximum was 1.69 (Item Primary/MR5), and all other non-extreme items had Infits smaller than 1.32. This is considered to reflect a good fit to the model. The person and item reliabilities were good (.97 and .95, respectively). After assessing some of the psychometric properties of the measures, it was necessary to look more closely at the variable map (Fig.1).

The Pre-operational/Single Representation stage presented two sets of item difficulties, i.e. items Pre-op/SR1, Pre-op/SR3, Pre-op/SR4, Pre-op/SR5 and Pre-op/SR8 were shown to be less difficult than items Pre-op/SR2, Pre-op/SR6 and Pre-op/SR7. This gap between items with the same predicted OHC suggests that there was a problem in designing these items. One hypothesis to explain this effect could be that they are more horizontally complex. The Preo-operational items are composed of four equal letters plus a different letter, requiring the participant only to discriminate a set of five simple stimuli, choosing the dissimilar one. The items Pre-op/SR2, Pre-op/SR6 and Pre-op/SR7 may have been more difficult because the letters provided as options, in each item, were closer in graphical terms. The item Pre-op/SR2, for example, was composed by four "O" and one "Q". The visual stimuli of both letters are graphically closer, differing by the little "dash" on the bottom of Q. Previous research has shown that the structure of cognitive processing is composed of cascade-like relations (Demetriou, Christou, Spanoudis, & Platsidou, 2002; Demetriou, Mouyi, & Spanoudis, 2008) between processes with increasing complexity, beginning with speed processing (the most basic component of the cognitive architecture), followed by perceptual discrimination, perceptual control, conceptual control, short-term memory, working memory

and, finally, reasoning processes. According to Demetriou, Mouyi and Spanoudis (2008), perceptual discrimination "reflects sheer speed of processing together with the processes required to discriminate between two simple stimuli and identify the target one" (p. 439). So, when comparing different stimuli, those whose difference are based on small tiny cues (e.g. the little dash of letter Q), demand a higher perceptual discrimination than those having more cues (e.g. comparing "A" with "E"). Thus, Pre-op/SR2, Pre-op/SR6 and Pre-op/SR7 are more *horizontally* complex than the other four Pre-operational items, because they demand a slight higher level of perceptual discrimination. In sum, it seems that in items from the Pre-operational order it is important to control as much as possible the perceptual discrimination required for the item or task, in order to avoid interference from the standpoint of horizontal complexity.

The next order's items also present two set of difficulties. The items Prim/RM1, Prim/RM2, Prim/RM3 and Prim/RM4 were the easiest items of the Primary stage, probably because they were constructed according to the Rule 1, i.e. four options with no jump between the pair of letters, and one option jumping one letter. The other four Primary items where constructed according to the Rule 2, which states a jump of one letter between each pair of letters (4 options), and one option jumping two letters. Our hypothesis is that when dealing with items constructed according to Rule 2, the participants needed to store and deal with more information in Working Memory (Demetriou et al., 2002, 2008; Pascual-Leone, 1984), which could horizontally increase the complexity of the task. A similar effect also seems to occur with the next order's items. Note the items Conc/RS5, Conc/RS6, Conc/RS7 and Conc/RS8, which are the most difficult concrete items, have a mean difference of .92 logits from the Conc/RS1, Conc/RS2, Conc/RS3 and Conc/RS4. This might be because the most difficult items have a rule which involves one more bit of information, being more horizontally complex than the items Conc/RS1, Conc/RS2, Conc/RS3 and Conc/RS4. Originally, we varied some of the rules somewhat in order to make the task less boring, and to avoid possible fatigue from the repetition of procedures employed to answer an item or task. However, our result suggests that changing some items' rules within a certain OHC can compromise the quality of the stage identification. It seems that a good strategy for developmental test construction is trying always to elaborate items with the same rule within a single OHC.

The items from the Abstract, Formal and Systematic orders, on the other hand, are forming groups, or clusters, reflecting the fact that items within each are of the same hierarchical complexity (and are therefore grouped together), and items across each order are appropriately separated. The Abstract items, however, are not well separated from the Concretes items. It can be speculated that the way the tables of the Abstract order were constructed, having eight code rows, each beginning with an alphabetic letter followed by a Greek letter, decreases the difficulty of the items. The options of the items are all organized and well structured, and this organization seems to work as a support for the respondents.

41	AδBηFπAμ	CηEπBλDδ	AδFπAμDδ	BπFλCδFη	BδEλAπFμ

Figure 12. Example: Item 41, Systematic/AS

Table 3. One-sample tests of mean item difficulties for different OHC's

| | Test value = 0 | | | | | | | | |
| | | | | | | 95% CI | | Effect size | |
Stages	*t*	*DF*	Sig. (2-tailed)	*M*	*SD*	Lower	Upper	(*d'*)	(*r*)
Pre-op/SR and Primary/RM	10.36	7	0.00	3.61	0.99	2.79	4.43	7.83	.96
Primary/RM and Concrete/RS	22.94	7	0.00	3.42	0.42	3.06	3.77	17.34	.99
Concrete/RS and Abstract/SA	23.03	7	0.00	3.33	0.41	2.99	3.67	17.40	.99
Abstract/AS and Formal/AM	10.96	7	0.00	1.14	0.29	0.89	1.38	8.28	.97
Formal/AM and Systematic/AS.	4.78	7	0.00	0.88	0.52	0.44	1.31	3.61	.87

In spite of providing good indicators of the items' structure, and enabling the verification of visual clusters of items, the Rasch analysis did not provide information regarding the size of the gaps between adjacent OHC. The one-sample t-tests, calculated for this purpose, showed that the differences between adjusted difficulties of items from adjacent orders are statistically significant, with large effect sizes. This provides some additional evidence that helps support the existence of developmental stages of inductive reasoning. However, this result should be carefully interpreted, and future studies should employ a more balanced sample, from childhood to adulthood.

Study II: Refining the IRDT and investigating its construct/congruent validity

Study 2 aims to modify some items of the IRDT, based on the results from the first study, and, using Rasch analysis, assess its new scale structure, verifying whether the previously predicted orders and gaps, as well as the scale's reliability and unidimensionality.

Part I: Instrument improvement

From the results of Study I, we've modified some items of the IRDT. Basically, the modifications can be synthesized as follows. From the original eight Pre-operational items, those demanding high perceptual discrimination were excluded, due to close similarities and low graphical clues (such as Q and O, etc), except one. We left one item to verify whether it still has more difficulties than the other Pre-operational items. The others were all modified in order to obtain items with easily discriminative options, such as "R F F F F" (Item Pre-op/SR3) and "H H L H H" (Item Pre-op/SR8). At the Primary order we removed those items constructed based on Rule 2, in which the pair of letters jumps one letter of the alphabetic sequence, and replaced them with

items constructed based on Rule 1, i.e. with no jump in the letters' sequence, except for the option that is the exception and therefore is correctly supposed to be chosen by the participants because it does not follow the rule. Finally, the last change in the instrument occurred with the Abstract items, more precisely in the tables where the coordination of Concrete sequences are displayed. Instead of having a specific alphabetic letter in each row, and a specific Greek letter in each column, forming a code composed by two symbols for each cell that contains a coordination of two Concrete sequences, the table was modified to contain only one symbol (Greek letter) per cell. Moreover, the Abstract items are now formed by options that are spread throughout the table, so the participant needs to locate each one, and try to figure out which has a coordination rule that differs from the other 4 options. In the first version of the IRDT, the Abstract items' options were organized in each row. Also, the "plus" (+) symbol that mediated the coordination of the two Concrete sequences was taken out. The other two orders' items remained the same, since they demand the coordination of actions from the previous adjacent OHC. In sum, we've remodeled the items within each order, focusing on its vertical complexity. Our hypothesis is that this "*verticalization*" provides a better stage identification, with visual clusters of items and gaps between adjacent OHC more clearly defined.
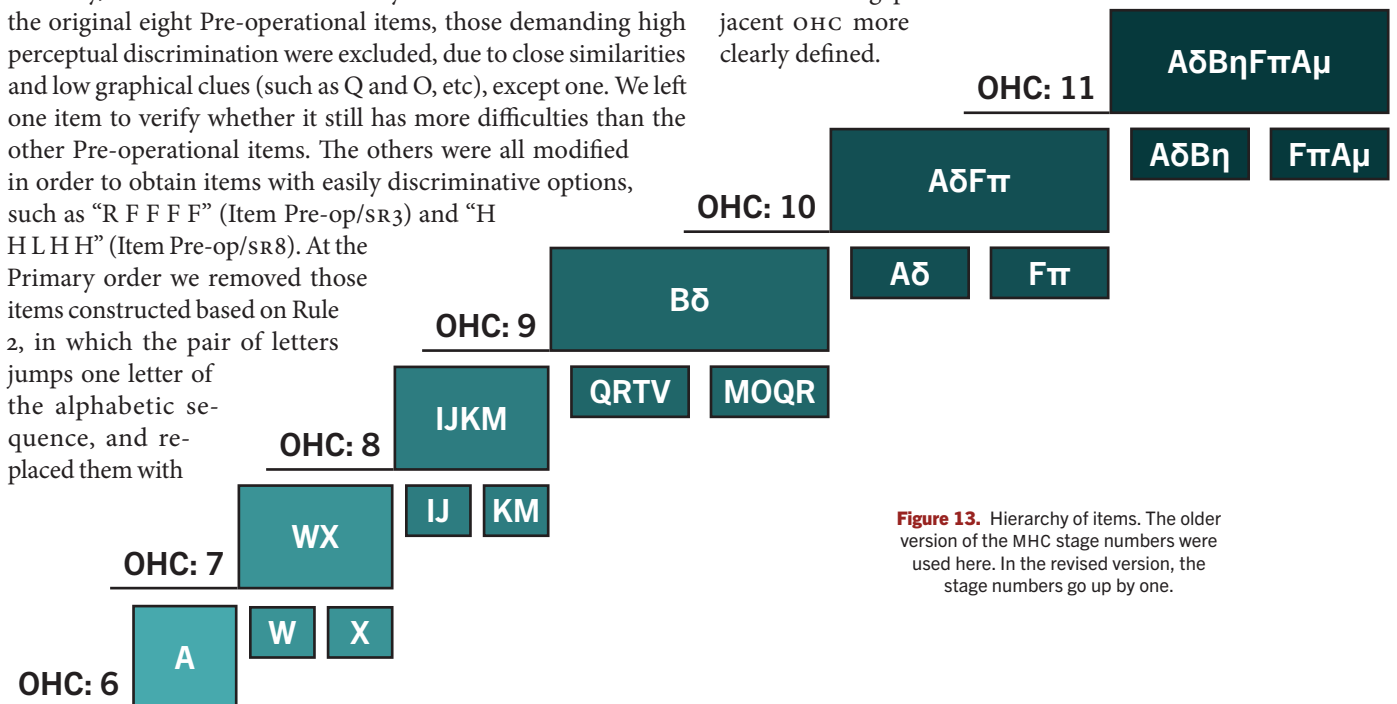
Figure 13. Hierarchy of items. The older version of the MHC stage numbers were used here. In the revised version, the stage numbers go up by one.
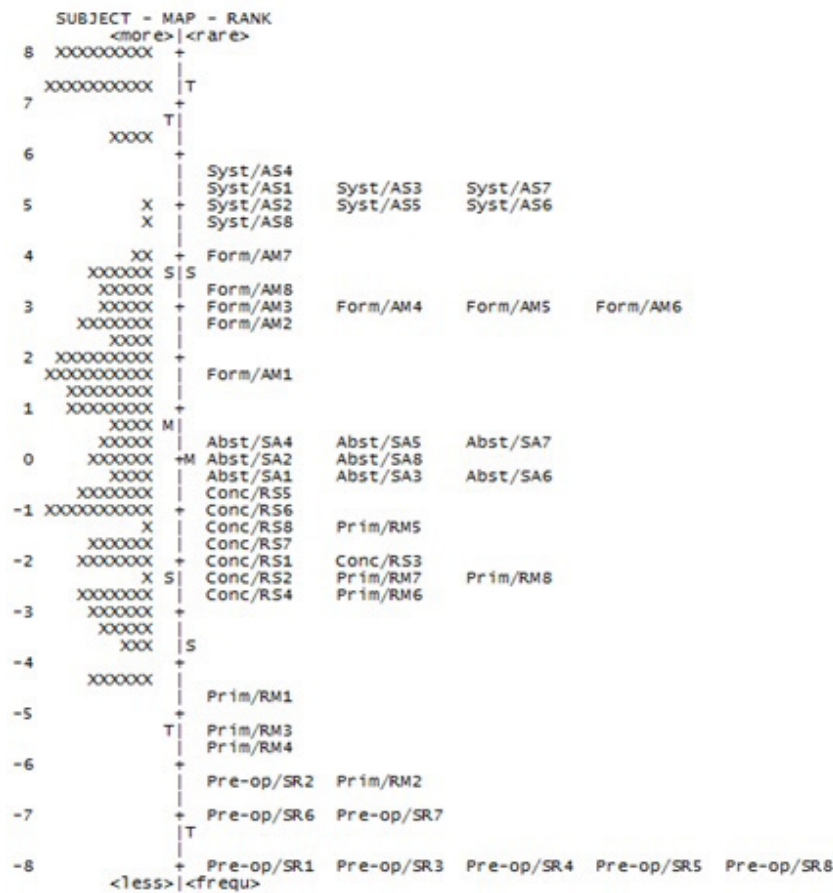
Figure 14. Variable Map showing the IRDT's items

» METHOD

Participants

In Study 2, the revised IRDT were administered to a convenience sample composed of 188 Brazilian people (42.3% men, 57.7% women) aged between 6 to 65 years (M = 21.45, SD = 14.31). The sample, again, was intentionally broad and had a distribution of 34.4% from 6 to 12 years, 13.4% from 13 to 15 years, 7.5% from 16 to 21 years, and 44.6% older than 21 years. All the participants were from the city of Belo Horizonte, state of Minas Gerais.

Procedure

The data were collect by the first author and by twenty five Psychology undergraduate students, enrolled in a second semester Cognitive Development class, who were trained to administer the instrument properly. The author first administered the instrument to the undergraduate students (and those which data are actually being used in this analysis). Each undergraduate student had to administer the IRDT to different people from 6 to 65 years old. Participation was voluntary. The potential participants had the purpose of the study explained to them. They were informed that their answers would be kept confidential, and that all procedures guaranteeing the privacy of their results would be adopted. They signed a inform consent, according to the guidelines of the Ethical Committee of the Universidade Federal de Minas Gerais, Brazil.

Data analysis

The same data analytic process presented in Study 1 was adopted here. To assess the new scale structure of the IRDT, verifying if it presents the predicted orders and gaps, as well as its reliability and unidimensionality, we've employed the dichotomous Rasch model. To verify if the differences between the mean difficulty of items from order k and the mean difficulty of items from order $k+1$ are statistically significant, the one-sample t-test is used, with 95% confidence interval. The effect size is calculated using Cohen's d.

» RESULTS

The Rasch dichotomous model (Andrich, 1988; Rasch, 1960) was calculated using the software Winsteps (Linacre, 1999, 2011). From 48 items, only one was correctly responded to by all participants (Pre-op/SR8). The reliability for the full scale was .99, and its Infit mean was .94 (SD = .22; Max = 1.46; Min = .56). The person reliability was .95, which is estimated to indicate the degree to which a person's response pattern conforms to the difficulty structure of the measure (Hibbard, Collins, Mahoney & Baker, 2009). The principal contrast showed that the raw variance explained by measures (modeled) was 74.8%, and that the unexplained variance in the first contrast (modeled) was 12.9%, suggesting that the instrument can be thought of as unidimensional, even though the variance explained by the first contrast is higher than 10%. We argue that the variance explained by measures (modeled) is high enough to sustain its unidimensionality.

The variable map (Figure 2) illustrates the scale for the 48 items of the IRDT with item difficulties (on the right) and person (student) measures (on the left) calibrated on the same scale. It's visually possible to identify clear item clusters for almost all the orders, with a gap between them. However, two formal items, Form/AM6 and Form/AM8 had their scaled difficulties closer to the Systematic items, and one additional formal item, Form/AM3, had its scaled difficulty closer to the Abstract items. The only other difficulties were with the Pre-operational items, which were very spread out, but were nevertheless separated from the Primary items. Regarding the relative position of person (left) and item (right), the variable map shows the IRDT was an easy test for 28 participants (*Mean ability* = 7.86, SD = 0.87). The whole-sample mean ability was 1.15 with standard deviation of 3.40 logits (see Figure 15).

The one-sample t-test, with 95% confidence interval, shows that the comparisons between Pre-operational and Primary, Primary and Concrete, Concrete and Abstract, Abstract and Formal, and between Formal and Systematic were significant. Moreover, the effect size d' and r were large (see Table 3).

» DISCUSSION

The evidence shows that modifying the IRDT, in order to eliminate some sources of horizontal complexity, produced an item structure closer to what was expected when constructing an instrument ac-

cording to the MHC and using the strategies presented in the introduction (see Figure 1). In each OHC, the items are grouped forming a visual cluster, and presenting a gap in relation to the adjacent orders. Two Formal items had difficulties higher than expected (Form/AM6 and Form/AM8) and one was less difficult than predicted. However, this small deviation does not interfere with the spacing of its Rasch scores in relation to the adjacent orders of hierarchical complexity. The Pre-operational items have its scaled difficulties somewhat scattered through the less difficult end of the scale, an unexpected result to some extent, since the items were modified to contain stimuli that were expected to be easily discriminated (having many graphical clues). However, it can be speculated that the differences in difficulty of these items are due to factors other than the nature of each stimulus' contribution to the increase in its horizontal complexity. In any case, the item Pre-op/SR4 presents a difficulty at least 1.26 logits higher than the other Pre-operational items. This result was expected, since the Pre-op/SR4 ("U U V U U") is the same in both versions of the IRDT, and presents options graphically close to each other, demanding a higher amount of perceptual discrimination.

Regarding the data's fit to the model, the modified version of the IRDT produced a better Infit mean of the items (.94), representing an increase of .06 over the items' Infit of the first version (.88). The percentage of variance explained by the measures also increased from 70.6 with the previous version to 74.8 with the new one. It can be speculated that when we eliminated part of the horizontal complexity of the items, the amount of variance explained by the unidimensional measure increased. So, the "verticalization" process seems to contribute to the measure, not only in terms of the theory behind the items, i.e. the Model of Hierarchical Complexity, and by consequence the expected item structure, but also in terms of the adjustment of the items to the model and to the amount of variance explained.

Now that the item structure is closer to the expected (Figure 1), and the items' fits are more adequate, it seems to be relevant to coordinate the Rasch metrics and the Orders of Hierarchical Complexity in a mathematical fashion, to obtain a score representing stage of performance. There is no direct way to obtain a person score that represents stage of performance from the estimates obtained through the Rasch Dichotomous model. This seems to be a dilemma, mainly because there is a difference in formal measurement theory terms between the OHC and the Rasch scores. The former is an analytic measure represented in an ordinal scale, while the latter are an empirical conjoint-interval measure. But, there's a way to calculate stage of performance from the Rasch estimates. It can be calculated only because the items have the properties previously expected, i.e. they form clusters or groups within each OHC, present significant gaps with higher effect size between adjacent orders, and have adequate fit to the Rasch model. So, meeting these conditions, one can apply the below formula:
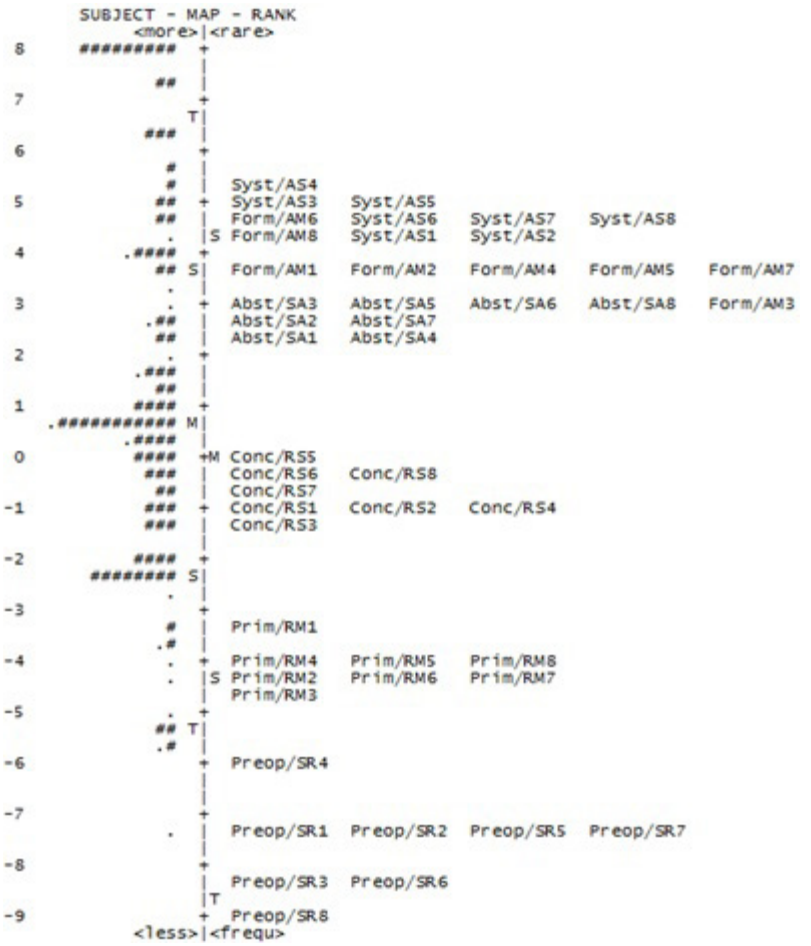
Figure 15. Variable Map showing the IRDT 2nd version's items

$$\varphi_j = \frac{\beta_j - \overline{X}_k}{\overline{X}_{k+1} - \overline{X}_k} + OHC_k \qquad (3)$$

where φ_j is the stage of performance of person j, β is the Rasch score of that person, $\overline{X}_k$ is the mean difficulty of items on order k, $\overline{X}_{k+1}$ is the mean difficulty of items on the next adjacent order, and OHC_k is the number that represents the order of hierarchical complexity k. For computing the stage scores of people whose ability lies on the highest order measured, one needs to leave the denominator as $\overline{X}_k$. After computing the stage of performance for each person, it is possible to verify how well the stage scores regress on the order of hierarchical complexity of the items. Figure 16 shows the linear regression. As can be seen, the Order of Hierarchical Complexity of an item predicted the mean performance on that item with an R^2 of 0.97 (see Figure 16).

» CONCLUSION

In line with previous researches (Bond & Fox, 2001; Commons et al., 2008; Dawson, 2000, 2002; Dawson, Xie, & Wilson, 2003; Dawson-Tunik, 2004; Dawson-Tunik, Commons, Wilson & Fischer, 2005), the current study adds supportive evidence for developmental stages using modern quantitative methods and a specific test design provided by the model of Hierarchical Complexity and by
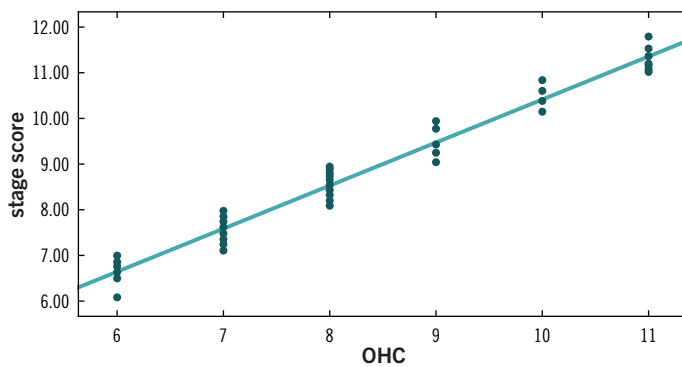
**Figure 16.** Regression of Stage Scores on Order of Hierarchical Complexity. The older version of the MHC stage numbers was used here. In the revised version, the stage numbers go up by one

the Dynamic Skill Theory. This study also adds a new group of instruments with extremely high r's between the order of hierarchical complexity used to predict the difficulty and the obtained difficulty. The difference between study 1 and 2 also shows the psychometric usefulness of constructing items with low horizontal complexity (number of actions) when what one is interested in is hierarchical complexity. Also of great importance, is that these instruments test all the way down to the preoperational stage and go up through the systematic stage. It would be easy to make a metasystematic version by asking people to compare the degree of similarity between systems from the systematic order—dissimilar, similar. Future studies should include higher stages and apply latent class analysis (Bond & Fox, 2001; Dawson-Tunik et. al., 2010; Demetriou & Kyriakides, 2006) of item's difficulties, since the use of different quantitative techniques can benefit stage evidences' strength. ■

REFERENCES

Andrich, D. (1988). *Rasch models for measurement*. Sage series on quantitative applications in the Social Sciences, Beverly Hills.

Andrich, D. (2002). Understanding Rasch measurement: Understanding resistance to the data-model relationship in Rasch's paradigm: A reflection for the next generation. *Journal of Applied Measurement*, 3(3), 325-359.

Andrich, D. (2004). Controversy and the Rasch model: a paradigm of incompatable paradigms. *Medical Care, 42*(1), 7-16.

Armon, C. (1984). Ideals of the good life and moral judgment: Ethical reasoning across the lifespan. In M. Commons & F. Richards & C. Armon (Eds.), *Beyond formal operations: Late adolescent and adult cognitive development, Vol 1*. (pp. 357-380). New York: Praeger.

Armon, C. & Dawson, T. L. (1997). Developmental trajectories in moral reasoning across the life-span. *Journal of Moral Education, 26*(4), 433-453.

Baillargeon, R. (1987). Object permanence in 3 1/2- and 4 1/2-month-old infants. *Developmental Psychology, 23* (5), 655-664.

Bart, W. (1971). The effect of interest on horizontal decaláge at the stage of formal operations. *Journal of Psychology: Interdisciplinary and Applied*, 78(2), 141-150. doi:10.1080/00223980.1971.9916897

Bernholt, S., Parchmann, I. & Commons, M. (March, 2008). *Hierarchical Complexity Applied to the Domain of Chemistry: An Educational Research and Modeling Approach*. Paper presented at the Society for Research in Adult Development, New York, NY.

Bidell, T. R., & Fischer, K. W. (1992). Beyond the stage debate: Action, structure, and variability in Piagetian theory and research. In R. J. Sternberg, C. A. Berg, R. J. Sternberg, C. A. Berg (Eds.), *Intellectual development* (pp. 100-140). New York, NY US: Cambridge University Press.

Bock, R.D., & Jones, L.V. (1968). *The Measurement and Prediction of Judgment and Choice*. San Francisco: Holden Day.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences (2nd ed.)*. Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Bowman, A. K. (1996a). *The relationship between organizational work practices and employee performance: Through the lens of adult development*. Unpublished doctoral dissertation. The Fielding Institute, Santa Barbara, CA.

Bowman, A. K. (1996b). Examples of task and relationship 4b, 5a, 5b statements for task performance, atmosphere, and preferred atmosphere. In M. L. Commons, E. A. Goodheart, T. L. Dawson, P. M. Miller, & D. L. Danaher, (Eds.) *The general stage scoring system*

(GSSS). Presented at the Society for Research in Adult Development, Amherst, MA.

Broughton, J.M. (1984). Not beyond formal operations, but beyond Piaget. In M. Commons, F.A. Richards, and C. Armon (Eds.), *Beyond formal operations: Late adolescent and adult cognitive development,* Vol 1, (pp. 395-411). New York: Praeger.

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

Chapman, M., & Lindenberger, U. (1988). Functions, operations, and decaláge in the development of transitivity. *Developmental Psychology*, 24(4), 542-551. doi:10.1037/0012-1649.24.4.542

Chazan, S. (1972, May). Horizontal decaláge in the concept of object permanence as a correlate of dimensions of maternal care. *Dissertation Abstracts International*, 32. 6613.

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 issue scoring manual*. New York: Cambridge University Press.

Commons, M. L. (2008). Introduction to the model of hierarchical complexity and its relationship to postformal action . *World Futures*, 64(5-7), 305-320. doi:10.1080/02604020802301105

Commons, M.L., Gane-McCalla, R., Barker, C.D., & Li, E.Y (2014). The Model of Hierarchical Complexity as a Measurement System. *Behavioral Development Bulletin, 19*(3).

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.

Commons, M. L., Lee, P., Gutheil, T. G., Goldman, M., Rubin, E. & Appelbaum, P. S. (1995). Moral stage of reasoning and the misperceived "duty" to report past crimes (misprision). *International Journal of Law and Psychiatry, 18*(4), 415-424. doi:10.1016/0160-2527(95)00017-C

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

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*(4), 323-340. doi:10.1016/0273-2297(90)90017-X

Commons, M. L., Rodriguez, J. A. (1993). The development of hierarchically complex equivalence classes. *Psychological Record, 43*(4), 667-697.

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*(6), 430-435.

Commons, M. L., & Richards, F. A. (1984a). Applying the general stage model. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), *Beyond formal operations. Late adolescent and adult cognitive development: Late adolescent and adult cognitive development,* Vol 1. (pp. 141–157). NY: Praeger.

Commons, M. L., Richards, F. A., & Kuhn, D. (1982). Systematic and metasystematic reasoning: A case for a level of reasoning beyond Piaget's formal operations. *Child Development, 53*(4), 1058-1069. doi:10.1111/1467-8624.ep8587686

Commons, M., Goodheart, E., Pekker, A., Dawson, T., Draney, K., & Adams, K. (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.

Commons, M., Trudeau, E., Stein, S., Richards, F., & Krause, S. R. (1998). Hierarchical complexity of tasks shows the existence of developmental stages. *Developmental Review*, *18*(3), 237-278. doi:10.1006/drev.1998.0467

Cook-Greuter, S. R. (1990). Maps for living: Ego-development theory from symbiosis to conscious universal embeddedness. In M. L. Commons, J. D. Sinnott, F. A. Richards, & C. Armon (Eds.). *Adult Development: Vol. 2, Comparisons and applications of adolescent and adult developmental models* (pp. 79-104). New York: Praeger.

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

Dawson, T. L. (2001). Layers of structure: A comparison of two approaches to developmental assessment. *Genetic Epistemologist, 29* (4)*,* 1-10*.*

Dawson, T. L. (2002). New tools, new insights: Kohlberg's moral reasoning stages revisited. *International Journal of Behavioral Development, 26*(2)*,* 154-166. doi:10.1080/01650250042000645

Dawson, T. L. (2003). A stage is a stage is a stage: A direct comparison of two scoring systems. *Journal of Genetic Psychology, 164*(3), 335-364.

Dawson, T. L. (2004). Assessing intellectual development: Three approaches, one sequence. *Journal of Adult Development, 11*(2)*,* 71-85*.*

Dawson, T. L. (2006). Stage-like patterns in the development of conceptions of energy. In X. Liu & W. Boone (Eds.), *Applications of Rasch measurement in science education* (pp. 111-136). Maple Grove, MN: JAM Press.

Dawson, T. L., & Wilson, M. (2004). The LAAS: A computerized developmental scoring system for small- and large-scale assessments. *Educational Assessment, 9*(3/4), 153-191.

Dawson, T. L., Xie, Y., & Wilson, M. (2003). Domain-general and domain-specific developmental assessments: Do they measure the same thing? *Cognitive Development*, *18*(1)*,* 61-78. doi:10.1016/S0885-2014(02)00162-4

Dawson, T., Goodheart, E., Draney, K., Wilson, M., & Commons, M. (2010). Concrete, abstract, formal, and systematic operations as observed in a 'Piagetian' balance-beam task series. *Journal of Applied Measurement*, *11*(1), 11-23.

Dawson-Tunik, T. L. (2004). "A good education is. . ." The development of evaluative thought across the life-span. *Genetic, Social, and General Psychology Monographs, 130*(1)*,* 4-112*.*

Dawson-Tunik, T. L., Commons, M., Wilson, M., & Fischer, K. (2005). The shape of development. *The European Journal of Developmental Psychology, 2,* 163-196.

Demetriou, A., & Kyriakides, L. (2006). The functional and developmental organization of cognitive developmental sequences. *British Journal of Educational Psychology*, *76*(2), 209 242. doi:10.1348/000709905X43256

Demetriou, A., Christou, C., Spanoudis, G., & Platsidou, M. (2002). The development of mental processing: Efficiency, working memory, and thinking. *Monographs of the Society of Research in Child Development*, *67*(1), vii-154. doi:10.1111/1540-5834.671173

Demetriou, A., Efklides, A., Papadaki, M., Papantoniou,G., & Economou, A. (1993). Structure and development of causal experimental thought: From early adolescence to youth. *Developmental Psychology, 29*(3), 480-497.

Demetriou, A., Mouyi, A., & Spanoudis, G. (2008). Modelling the structure and development of g. *Intelligence*, *36*(5), 437-454. doi:10.1016/j.intell.2007.10.002

Embretson, S.E. and Reise, S. P. (2000). *Item response theory for psychologists.* London: Erlbaum.

Feldman, D.H. (2004). Piaget's stages: The unfinished symphony of cognitive development. *New Ideas in Psychology*, *22*(3), 175-231. doi:10.1016/j.newideapsych.2004.11.005

Fischer, K. W. (2008). Dynamic cycles of cognitive and brain development: Measuring growth in mind, brain, and education. In A. M. Battro, K. W. Fischer, P. J. Léna, A. M. Battro, K. W. Fischer, P.

J. Léna (Eds.), *The educated brain: Essays in neuroeducation* (pp. 127-150). New York, NY US: Cambridge University Press.

Fischer, K. W., & Bidell, T. R. (1998). Dynamic development of psychological structures in action and thought. In W. Damon, R. M. Lerner, W. Damon, R. M. Lerner (Eds.), *Handbook of child psychology: Volume 1: Theorectical models of human development (5th ed.)* (pp. 467-561). Hoboken, NJ US: John Wiley & Sons Inc.

Fischer, K. W., & Bidell, T. R. (2006). Dynamic development of action, thought, and emotion. In W. Damon & R. M. Lerner (Eds.), *Theoretical models of human development*. *Handbook of child psychology* (6th ed., Vol. 1, pp. 313-399). New York: Wiley.

Fischer, K. W., & Yan, Z. (2002a). The development of dynamic skill theory. In R. Lickliter & D. Lewkowicz (Eds.), *Conceptions of development: Lessons from the laboratory*. Hove, U.K.: Psychology Press.

Fischer, K. W., & Yan, Z. (2002b). Darwin's construction of the theory of evolution: Microdevelopment of explanations of variation and change of species. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition Processes in Development and Learning*. Cambridge, U.K.: Cambridge University Press.

Fischer, K.W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review, 87*(6), 477-531. doi:10.1037/0033-295X.87.6.477

Fischer, K.W. (1987). Relations between brain and cognitive development. *Child Development, 58*(3), 623-632.

Fischer, K.W., & Dawson, T. L. (2002). A new kind of developmental science: Using models to integrate theory and research. *Monographs of the Society for Research in Child Development*, *67* (1), 156—167. doi:10.1111/1540-5834.671184

Fischer, K.W., & Rose, S.P. (1994). Dynamic development of coordination of components in brain and behavior: A framework for theory and research. In G. Dawson & K.W. Fischer (Eds.). *Human behavior and the developing brain* (pp. 3-66). New York: Guilford Press.

Fischer, K.W., & Rose, S.P. (1999). Rulers, models, and nonlinear dynamics: measurement and method in developmental research. In G. Savelsbergh, H. van der Maas: Royal Netherlands Acadamy of Arts and Sciences.

Fischer, K.W., & Silvern, L. (1985). Stages and individual differences in cognitive development. *Annual Review of Psychology, 36*, 613-648. doi:10.1146/annurev.ps.36.020185.003145

Fischer, K.W., Hand, H.H., & Russell, S. (1984). The development of abstractions in adolescence and adulthood. In M. Commons, F.A.

Richards, and C. Armon (Eds.), *Beyond formal operations: Late adolescent and adult cognitive development,* Vol 1, (pp. 43-73). New York: Praeger.

Fischer, K.W., Kenny, S.L., & Pipp, S.L. (1990). How cognitive processes and environmental conditions organize discontinuities in the development of abstractions. In C.N. Alexander, E.J. Langer, & R.M. Oetzel (Eds.), *Higher stages of development,* (pp. 162-187). New York: Oxford University Press.

Flavell, John H. (1963). *The Developmental Psychology of Jean Piaget*. Princeton, NJ: Van Nostrand.

Glas, C.A. (2007). *Multivariate and Mixture Distribution Rasch Models*. New York: Springer-Verlag.

Gomes, C. M. A. (2010). Estrutura fatorial da Bateria de Fatores Cognitivos de Alta-ordem (BAFACALO). *Avaliação Psicológica, 9*, 449-459.

Gomes, C.M.A. & Golino, H.F. (2009). Estudo exploratório sobre o Teste de Desenvolvimento do Raciocinio Indutivo (TDRI). In D. Colinvaux. *Anais do VII Congresso Brasileiro de Psicologia do Desenvolvimento: Desenvolvimento e Direitos Humananos.* (pp. 77-79). Rio de Janeiro: UERJ. Available in http://www.abpd.psc.br/files/congressosAnteriores/AnaisVIICBPD.pdf

Gomes, C. M. A. & BORGES, O. N. (2009). Qualidades Psicométricas do Conjunto de Testes de Inteligência Fluida. *Avaliação Psicológica*, *8*, 17-32.

Goodheart, E. A., Dawson, T. L. (1996, June). *A Rasch Analysis of Developmental Data from The Laundry Problem Task Series.* Poster presented at the 11th Annual Adult Development Symposium, Boston, MA.

Goodheart, E. A., Dawson, T. L., Draney, K., Commons, M. L. (1997, March). *A Saltus Analysis of Developmental Data from The Laundry Problem Task Series.* Poster presented at IOMW9, Chicago, IL.

Halford, G.S. (1989). Reflexions on 25 years of Piagetian cognitive developmental psychology, 1963–1988. *Human Development, 32,* 325-357.

Hambleton, R. K. & Jones, R. W. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice, 12*(2), 38-47. doi:10.1111/j.1745-3992.1993.tb00543.x

Hambleton, R. K. (2000). Emergence of item response modeling in instrument development and data analysis. *Medical Care*, 38(Suppl9), II60-II65. doi:10.1097/00005650-200009002-00009

Hartelman, P. A., van der Maas, H. J., & Molenaar, P. M. (1998). Detecting and modeling developmental transitions. *British Journal of Developmental Psychology*, 16(Pt 1), 97-122. doi:10.1111/j.2044-835X.1998.tb00751.x

Hibbard, J., Collins, P., Mahoney, E., & Baker, L. (2010). The development and testing of a measure assessing clinician beliefs about patient self-management. *Health Expectations: An International Journal of Public Participation in Health Care & Health Policy*, *13*(1), 65 72.

Jackson, E., Campos, J.J. & Fischer, K.W. (1978). The question of decaláge between object permanence and person permanence. *Developmental Psychology, 14*(1). 1-10. doi:10.1037/0012-1649.14.1.1

Jamison, W. (1977). Developmental inter-relationships among concrete operational tasks: An investigation of Piaget's stage concept. *Journal of Experimental Child Psychology*, *24*(2), 235–253.

Joaquim, C. J. (2014). Developmental stage of performance in reasoning about bullying in school age youth. *Behavioral Development Bulletin*, *19*(2), 81-99.

Kallio, E., & Helkaman, K (1991). Formal operations and postformal reasoning: A replication. *Scandinavian Journal of Psychology*, 32(1), 18-21. doi:10.1111/j.1467-9450.1991.tb00848.x

Kitchener, K. S. & Fischer, K. W. (1990). A skill approach to the development of reflective thinking. In D. Kuhn (Ed.), *Developmental perspectives on teaching and learning thinking skills. Contributions to Human Development*: Vol. 21, (pp. 48-62). Switzerland: Thur AG Offsettdruck

Kitchener, K. S., & King, P. M. (1990). Reflective judgement: Ten years of research. In M. L. Commons, C. Armon, L. Kohlberg, F. A. Richards, T. A. Grotzer, & J. D. Sinnott (Eds.), *Beyond formal operations: Vol. 2. Models and methods in the study of adolescent and adult thought* (pp. 63-78). New York: Praeger.

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

Kreitler, S., & Kreitler, H. (1989). Horizontal Decaláge: A problem and its solution. *Cognitive Development, 4*(1), 89-119. doi:10.1016/0885-2014(89)90006-3

Lautrey, J., de Ribaupierre, A., & Rieben, L. (1985). Intraindividual variability on the development of concrete operations: Relations between logical and infralogical operations. *Genetic, Social, and General Psychology Monographs*, *111*(2), 167-192.

Linacre, J. M. (1999). Investigating rating scale category utility. *Journal of Outcome Measurement, 32*(2), 103-122.

Linacre J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions, 16* (**2**)**, 878**

Linacre J. M. (2011). WINSTEPS. Rasch measurement computer program, Winsteps.com, Chicago.

Lourenço, O. (1998). Além de Piaget? Sim, mas Primeiro Além da Sua Interpretação Padrão! *Análise Psicológica*, 16(4), 521-552.

Lovell, C. W. (2002). Development and disequilibration: Predicting counselor trainee gain and loss scores on the Supervisee Levels Questionnaire. *Journal of Adult Development, 9*(3), 235-240. doi:10.1023/A:1016060328253

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

Marshall, P. E. (2009). *Positive psychology and constructivist developmental psychology: A theoretical* enquiry into how a developmental stage conception might provide further insights into specific *areas of positive psychology.* Unpublished Msc dissertation. University of East London, School of Psychology. Retrieved from http://devtestservice.org/about/articles.html

Martorano, S. (1977). A developmental analysis of performance on Piaget's formal operations tasks. *Developmental Psychology*, *13*(6), 666-672. doi:10.1037/0012-1649.13.6.666

Miller, J. G., Bett, E. S., Ost, C. M., Commons, M.L., Day, J. M., Robinett, T. L., Ross, S. N., Marchand, H. & Lins, M. da Costa (2008, June). *Finding the Relationships Among Moral Development Measures Using the Model of Hierarchical Complexity and Rasch Analysis*. Paper presented at Jean Piaget Society, Quebec City, Quebec, Canada.

Miller, J. G., Harrigan, W. J., Commons, M. L. & Commons-Miller, N. H. K. (2008, November). *An Analysis of Causing Religious Belief and Atheism Instruments and Hierarchical Complexity*. Paper presented at the Association for Moral Education, Notre Dame University, South Bend, Indiana.

Miller, P. (2002). *Theories of developmental psychology (4th ed.)*. New York, NY US: Worth Publishers.

Miller, P. M., &. Lee, S. T. (June, 2000). *Stages and transitions in child and adult narratives about losses of attachment objects.* Paper presented at the Jean Piaget Society. Montreal, Québec, Canada.

Morra, S., Gobbo, C., Marini, Z., & Sheese, R. (2008). *Cognitive development: Neo Piagetian perspectives*. New York, NY: Taylor & Francis Group/Lawrence Erlbaum Associates.'

Murray, F. (1969). Conservation in self and object. *Psychological Reports*, *25*(3), 941-942. doi:10.2466/pr0.1969.25.3.941

Murray, F., & Holm, J. (1982). The absence of lag in conservation of discontinuous and continuous materials. *Journal of Genetic Psychology*, *141*(2), 213-217.

Nummedal, S. (1971). The existence of the substance-weight-volume decaláge. *Dissertation Abstracts International*, 31, 3347.

Panayides, P., Robinson, C., & Tymms, P. (2010). The assessment revolution that has passed England by: Rasch measurement, *British Educational Research Journal*, *36*(4), 611 626. doi:10.1080/01411920903018182

Pascual-Leone, J. (1984). Attentional, Dialectic, and Mental Effort: Toward an Organismic Theory of Life Stages, In Michael L. Commons, Francis A. Richards and Cheryl Armon (Eds.), *Beyond formal operations. Late adolescent and adult cognitive development,* Vol 1. (pp. 182-215). New York: Praeger.

Peeters, M.J. & Stone, G.E. (2009). An Instrument to Objectively Measure Pharmacist Professionalism as an Outcome: A Pilot Study. *The Canadian Journal of Hospital Pharmacy, 62*(3), 209-216.

Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years*. New York: Holt, Rinehart, & Winston.

Rasch, G. (1960/1993). *Probabilistic models for some intelligence and attainment tests.* (Copenhagen, Danish Institute for Educational Research). Expanded edition (1980) with foreword and afterword by B.D. Wright, (1980).Chicago: MESA Press.

Richardson, A. M. & Commons, M. L. (2008, July). *Accounting for Stage of Development on Mathematical and Physical Science Tasks*. Paper presented at the Society for Mathematical Psychology, Washington D.C.

Rijmen, F., De Boeck, P., & Van der Mass, H. J. (2005). An IRT Model with a Parameter Driven Process for Change. *Psychometrika*, *70*(4), 651-699. doi:10.1007/s11336-002-1047-z

Roberge, J. (1976). Developmental analyses of two formal operational structures: Combinatorial thinking and conditional reasoning. *Developmental Psychology*, *12*(6), 563 564. doi:10.1037/0012-1649.12.6.563

Rose, S. P., & Fischer, K. W. (1998). Models and rulers in dynamical development. *British Journal of Developmental Psychology*, *16*(1), 123-131. doi:10.1111/j.2044-835X.1998.tb00752.x

Rosnow, R. L., & Rosenthal, R. (1996). Computing contrasts, effect sizes, and counternulls on other people's published data: General procedures for research consumers. *Pyschological Methods, 1*(4), 331-340. doi:10.1037/1082-989X.1.4.331

Salzberger, T. (2011). 'The role of the unit in physics and psychometrics' by Stephen Humphry—One small step for the Rasch model, but possibly one giant leap for measurement in the social sciences. *Measurement: Interdisciplinary Research and Perspectives, 9*(1), 59-61. doi:10.1080/15366367.2011.558792

Scardamalia, M. (1977). Information processing capacity and the problem of horizontal Decaláge : A demonstration using combinatorial reasoning tasks. *Child Development*, *48*(1), 28-37.

Schwartz, M. S., & Fischer, K. W. (2005). Building general knowledge and skill: Cognition and microdevelopment in science learning. In A.

Demetriou & A. Raftopoulos (Eds.), *Cognitive developmental change: Theories, models, and measurement*. Cambridge, U.K.: Cambridge University Press.

Siegler, R., & Crowley, K. (1991). The gospel of Jean Piaget, according to John Flavell. *PsycCRITIQUES*, *36*(10), 829-831. doi:10.1037/030229

Siegler, R.S, (1981). Developmental sequences within and between concepts. *Monograph of The Society for Research in Child Development*, *46*(2), 1-84. doi:10.2307/1165995

Smith, L. (2002). From epistemology to psychology in the development of knowledge. In T. Brown, L.

Smith, T. Brown, L. Smith (Eds.), *Reductionism and the development of knowledge* (pp. 201-228). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.

Sonnert, G., & Commons, M. L. (1994). Society and the highest stages of moral development. *Politics and the Individual, 4*(1), 31-55.

Stein, Z., & Hiekkinen, K. (2009). Metrics, models, and measurement in developmental psychology. *Integral Review, 5*(1), 4-24.

Stein, Z., Dawson, T., & Fischer, K. W. (2010). Redesigning testing: Operationalizing the new science of learning. In M. S. Khine & I. M. Saleh (Eds.), *New Science of Learning: Cognition, Computers and Collaboration in Education* (pp. 207−224). New York: Springer.

Van der Maas, H. L., & Molenaar, P. C. M. (1992). Stagewise cognitive development: An application of catastrophe theory. *Psychological Review,99*(3)*, 395−417. doi:10.1037/0033-295X.99.3.395

Van Geert, P. & Steenbeek, H. (2005). Explaining after by before: Basic aspects of a dynamic systems approach to the study of development. *Developmental Review*, 25, 408 442. doi:10.1016/j.dr.2005.10.003

Webb, R. (1974). Concrete and formal operations in very bright 6- to 11-year-olds. *Human Development*, *17*(4), 292-300. doi:10.1159/000271352

Wright, B. D. & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*(3), 370.

Wright, B.D. and Stone, M.H. (1979). *The measurement model. Best Test Design: Rasch Measurement*, Chicago: Mesa Press

Yan, Z., & Fischer, K. W. (2007). Pattern emergence and pattern transition in microdevelopmental variation: Evidence of complex dynamics of developmental processes. *Journal of Developmental Processes*, *2*(2), 39-62.

APPENDIX A

Description of the IRDT demands by OHC

Note. The older version of the MHC stage numbers was used here. In the revised version, the stage numbers go up by one.

OHC	Name	What they do	How they do	
6	preoperational	Make very simple logical inductions, from single stimulus.	Proceeds from the identification and analysis of a group of single (equal) letters to a conclusion about an individual letter.	Distinguish single categories from each other (e.g. equal letters vs. different letter) in order to make a logical conclusion.
7	primary	Simple logical induction, from coordinated stimulus.	Proceeds from the identification of the relation between two coordinated letters, to a conclusion about a specific coordinated pair of letters.	Maps relations between pair of stimuli, and compare a series of paired relations in order to make a logical conclusion.
8	concrete	Logical induction from a system of mapped stimulus.	Proceeds from the analysis of X pair of coordinated letters, forming a system of relations within a single option, to a conclusion about a specific coordination of X pair of letters.	Analyze a system of relations between stimuli, and compare the systems to make a logical conclusion.
9	abstract	Logical induction carried out through the comparison of single abstract, general, class of systems.	Proceeds from the identification and comparison of variables out of finite classes, to a conclusion about a specific variable.	Distinguish single, general, abstract variables, in order to make a logical conclusion.
10	formal	Logical induction from the coordinated abstract, general, class of systems.	Proceeds from the identification of the relation between two coordinated abstract variables, to a conclusion about a specific coordinated pair of variables.	Relationships are formed out of variables; mapping the relations to make a logical conclusion.
11	systematic	Logical induction from a system of mapped abstract, general, variables.	Proceeds from the analysis of X pair of coordinated abstract variables, forming a system of relations within a single option, to a conclusion about a specific coordination of X pair of abstract variables.	Analyze a system of relations between abstract, general variables, and compare the systems to make a logical conclusion.

APPENDIX B

Inductive Reasoning Developmental Test 2nd Version

		Pre-operational Items			
1	A	A	A	A	E
2	B	B	B	C	B
3	R	F	F	F	F
4	U	U	V	U	U
5	Q	Q	C	Q	Q
6	V	V	V	S	V
7	D	G	D	D	D
8	H	H	L	H	H

		Primary Items			
9	WX	KL	ST	PR	YZ
10	IJ	RT	CD	UV	MN
11	TU	HI	QR	JL	BC
12	PQ	NO	GI	CD	RS
13	XY	AB	TU	DF	OP
14	ST	IK	YZ	VW	EF
15	JK	DE	UV	HI	NP
16	GH	XZ	LM	RS	KL

		Concrete Items			
17	NOPR	IJKM	UVXY	MNOQ	QRSU
18	PQRT	LMNP	GHIK	VWXZ	KLNO
19	HIJL	TUWX	RSTV	OPQS	FGHJ
20	JKLN	BCDF	PQST	CDEG	STUW
21	OQST	DFHI	MOQR	EGHJ	TVXY
22	RTVW	ACEF	BDEG	CEGH	FHJK
23	IKMN	LNPQ	RTVW	JLMO	SUWX
24	GIKL	FHIK	PRTU	QSUV	CEGH

		Reference Table			
Ж	Ю	ф	ə	ħ
FGIKOQST	OPRTDFHI	IJLNPRSU	EFHJTVXY	RSUWNPRS
µ	π	σ	Љ	И
QRTVMOQR	STVXIKMN	KLNSUWX	CDFHNORS	GHJLPRTU
Ω	Σ	Δ	¥	Ħ
LMOQEGIJ	BCEGJLNO	MNPRGIKL	JKMOUWYZ	KLNPDEHI
Θ	Ξ	Π	Ψ	A
UVXKLNP	QSTVACEF	OQRTBDFG	FHIKRTVW	HJKMGIKL
œ	Ŧ	ø	β	δ
OPQTCEGH	JLMOPRTU	UWXZQSUV	CEFHNOPS	HJKMDFHI
Ђ	љ	ε	ζ	λ
KMNPGIKL	EGHJGHIL	QSTVMOQR	TVYKMOP	DFGISUWX
Щ	‡	þ	ʃ	ґ
CDGHUVWZ	KLOPEFGJ	CDGHTUVY	LMPQDEGI	QRUVMNOR
Б	ц	Ъ	ł	ŋ
TUXYIJKN	OPSTFGK	HILMNOPS	ABEFBCDG	UVYZJKMO

		Abstract Items			
25	Ж	Ю	Ф	Э	ђ
26	μ	π	σ	Љ	И
27	Ω	∑	∆	¥	Ħ
28	Θ	Ξ	Π	Ψ	α
29	œ	Ŧ	ǿ	β	δ
30	Ђ	љ	ε	Z	λ
31	Щ	‡	Þ	∫	Ґ
32	Б	Џ	Ъ	Ŧ	η

		Formal Items			
33	ЮЂ	∆œ	πδ	∑ε	ђμ
34	Жζ	Ю∆	¥α	Ωǿ	эΞ
35	ЮΩ	μλ	σŦ	ИΞ	ђΨ
36	ΨҐ	ǿ ɫ	Œλ	αЏ	Ŧþ
37	ЂЪ	δБ	λ‡	Ξ ε	Ξ Щ
38	λπ	α¥	ǿ μ	Ђσ	δ Π
39	ε œ	œЮ	Ψђ	Ξ Ω	ε И
40	Ŧэ	ПЖ	δ∑	ζ∆	ε Ψ

		Systematic Items			
41	Юσђэ	∆ œ И Ξ	ЮЂэ Ξ	μ λ∑ ε	π δ Жζ
42	¥ αИ Ξ	∆ ИŦ μ	σ Ψ ∆ Ŧ	π αэǿ	ђ Ψ ђ Ђ
43	σ α∑ Π	ђ ǿ Ω δ	Юζα∆	¥ δ Ω Ђ	μ Ξ∑ε
44	П Ж ǿ μ	ǿ Ω σ œ	δ∑Ξ Ω	λπ ε И	ЂσœЮ
45	α¥ Ŧ э	λ∆ ЂЖ	ǿ∑œэ	ε Ω Ξ ђ	ζπ μ α
46	Ψ ∑ μ α	ǿ ЮΨ Ω	П эζμ	Ŧ ђ ЂИ	α∆ ǿ ¥
47	Ψ Џ Ξ ɫ	Ŧ Ґδ Џ	λþα‡	αþσ α	ǿ ЩЂҐ
48	δ БαЏ	Ψ Ґ λ‡	ǿ ɫ Ξ Щ	Ŧ þЂЪ	ζЪ∑ П

BEHAVIORAL DEVELOPMENT BULLETIN

The defining issues test of moral judgment development

Stephen J. Thoma and Yangxue Dong

The University of Alabama

ABSTRACT

A measure of moral judgment development, the Defining Issues Test (DIT) is described and the supporting evidence for the measure is summarized. We address these questions: what does the DIT measure; how does the measure work, and how has the measure been validated? The psychometric properties of the DIT are also presented. We suggest that the current evidence supports the DIT as a reliable and valid measure of the characteristic ways adolescents and adults comprehend moral issues.

KEYWORDS: defining issues test, moral judgment, development

The Defining Issues Test (hereafter the DIT) was first developed in the early 1970s (Cooper, Coder, Masanz and Anderson, 1974). Originally the measure was described as a paper and pencil alternative to Lawrence Kohlberg's (1969) semi-structured interview measure of moral judgment development (Rest, 1979). As such, the primary focus of the measure was an assessment of the understanding and interpretation of moral issues. Consistent with the Kohlbergian model, Rest viewed moral judgment development as a social and cognitive construct that progressed from a self-focused view of moral issues, through a group-based moral perspective, to a reliance on post-conventional moral principles. Also consistent with Kohlberg, Rest viewed moral judgments as primarily cognitive and a primary factor in the understanding of moral actions and emotions. In short and during the 70s the DIT was viewed as a measure designed to test Kohlberg's developmental sequence and contribute to the development of moral judgment theory in adolescent and adult populations.

Although different in structure from Kohlberg's interview assessment, Rest borrowed the basic components of the Kohlberg approach. Similar to Kohlberg's moral judgment interview the DIT used stories to focus the participant on a moral dilemma. Many of these stories were originally used by Kohlberg (e.g., the story of Heinz and the drug). Furthermore, many of the items used on the DIT were based on Kohlberg interview data. However, unlike the Kohlberg interview where an individual must produce a response, the DIT is a recognition measure. On the DIT, participants are required to rate and then rank 12 short issue statements. These statements represent the defining features of the moral dilemma as viewed from each of Kohlberg's six-stages (Rest, 1979). Specifically, participants taking the DIT read the story and then decide what the protagonist ought to do (e.g., on the Heinz dilemma the choices are "steal the drug", "not steal" or "can't decide"). Following this action

choice, 12 items are presented and rated in terms of importance on a 5-point scale (from great importance to no importance). Once completed, the participant is asked to consider the 12 items as a set and then rank the four items that best describe their understanding of how the protagonist ought to solve the dilemma. This process is repeated for the remaining stories.

The primary index of moral judgment development is derived from the four items ranked as most important. Rest and his colleagues demonstrated that the DIT scores produced results that were consistent with theoretical expectations based on Kohlberg's model (e.g., Kohlberg, 1969). As described below, research using the DIT supported Kohlberg's claim that moral judgment is developmental and increases rapidly across high school and college years. Additionally, the scores produced by the DIT were able to distinguish groups of individuals who could reasonably be expected to differ on moral judgment development, were able to demonstrate that the measure was sensitive to educational interventions, and could related to moral actions and choices. Thus, Rest claimed, one could measure moral judgment development without having to interview individuals, interpret and score their verbal protocols.

It should be noted that in addition to the similarities between the DIT and Kohlberg's method, the DIT also shares some commonalities with measures derived from the Model of Hierarchical Complexity (MHC). Both approaches acknowledge their theoretical and methodological ties to Kohlberg and Piaget while modifying the assessment process and the definition of underlying constructs (e.g., Commons & Pekker, 2007). Additionally, both approaches yield measures that are claimed to be developmental and reflect a hierarchical integration of increasingly complex information. These similarities notwithstanding, there are some significant differences in both the measurement process and intended outcomes. As mentioned above, the DIT uses a rating and ranking task to identify the characteristic way the individual interprets moral situations. These estimates of

Author information: Steve Thoma, Educational Psychology, 310 Carmichael Hall, University of Alabama, Tuscaloosa, AL 35487-0231, http://www.ethicaldevelopment.ua.edu

moral judgments are described using a developmental model located solely within the moral domain. By contrast, the MHC uses multiple assessments (problem solving tasks, reasoning about vignettes and interviews) to generate an estimate of the individual's developmental stage, which encompasses moral judgments but is not limited to them. As a more general measure of development, the MHC emphasizes the participant's generalized ability to integrate information using estimates of performance on increasingly difficult tasks (Commons & Pekker, 2007). Thus, the MHC represents a measurement system that is not tied to any particular domain but can be used to inform our understanding of the moral domain.

What does the DIT measure?

The original interpretation of the DIT and what it measured reflected its association with Kohlberg's model. In this view, the DIT was a user-friendly methodological alternative to the interview method. However, this characterization no longer holds (Thoma, 2002; 2006). As the theoretical foundation of the measure evolved from Kohlberg's model to Rest's Four Component Model (Rest, 1983) the interpretation of what the DIT measures also changed. These changes are outlined below.

Changes related to the underlying developmental model. Early in the development of the DIT, Rest questioned Kohlberg's acceptance of a strong stage model of development in which individuals move from stage to stage one stage at a time. Instead, the DIT supported a developmental model that defines growth as a gradual shift from lower to more complex conceptions of social/moral cooperation. Furthermore, DIT researchers assume that at any given time there are multiple conceptions available to the individual. Thus, appropriate measurement strategies must assess not only which conceptions are available, but the most preferred system.

Additionally, in the 1990s, DIT researchers adopted a schema view of moral judgment development. A transition that signaled an abandonment of cognitive operations as the defining features of moral stages that was so central to Kohlberg's stage definitions (Kohlberg, 1984). This schema-based model represented moral development as a developmentally ordered set of schemas which define the network of knowledge that is organized around particular life events and exist to help individuals understand new information based on prior experiences (Rest, Narvaez, Bebeau, & Thoma, 1999). In this view moral schemas are contextual, automatic, and less reflective than Kohlberg's stages. Consistent with this view is a companion position suggesting that schemas may not be explicitly understood by the individual and may operate at the tacit level. Thus DIT researchers argue that the DIT is best viewed as a device for activating moral schema (Narvaez and Bock, 2002).

The schemas activated by the DIT are further claimed to be the most general and context-free system for interpreting moral situations. These schemas are labeled as "bedrock schemas" to distinguish the level of assessment provided by the DIT from more context depended interpretive systems. More specifically, the schemas measured by the DIT are viewed as a default system that is evoked when other, more automatic and context-specific, interpretive systems fail or provide incomplete or inconsistent information.

Micro vs. macro morality. It has been helpful to maintain a distinction between micro morality, or the morality of everyday exchanges, and macro-morality, or reasoning which focuses on society-wide considerations (e.g., Rest, Narvaez, Bebeau, & Thoma, 1999; Thoma, 2002). DIT researchers argue that the DIT assesses macro morality. In this view, what is assessed are default or bedrock schema that capture an individual's understanding of social cooperation in terms of justice and fairness within the context of law, the mechanisms of government and other social institutions. Although one can conceptually distinguish micro and macro morality, in practice one must assume that they overlap. However, DIT researchers claim that everyday morality is much more contextually dependent than macro morality and influenced by multiple interpretive systems that include but are not limited to the default system measured by the DIT.

Although DIT researchers make a distinction between macro and micro morality this does not imply that the impact of macro-morality on moral function is limited. Indeed, the significance of macro moral processes in adolescence and adulthood is often noted (e.g., Adelson, 1971; Torney-Purta, 1990). In fact, Rest and colleagues argue that the DIT measurement system assumes that the major developmental shifts during adolescence and beyond are the growing understanding of macro-moral conceptions of social cooperation in conventional and post-conventional terms (Rest, Narvaez, Bebeau, & Thoma, 1999).

How does the DIT work?

As mentioned above the DIT presents participants with a moral dilemma and then asks them to rate and rank 12 items for each dilemma. Each of the items raise particular issues that define the central features of the dilemma based on different moral schema considerations. These items do not present a complete rationale and interpretation of the dilemma but provide the gist of an explanation using a sentence fragment approach. The sentence fragment approach was adopted because early on in the development of the DIT it was noted that items which contained more detailed interpretations of the dilemmas yielded poor developmental indices in part because these items were prone to reinterpretation and idiosyncratic responding (Rest, 1979). By contrast, the use of sentence fragments are particularly well suited to trigger a schema because the fragment provides just enough information to suggest an interpretation, and the individual must fill in the necessary information to fully make sense of the item. Thus, DIT items which match the participant's preferred schema are rated as important and are candidates for being ranked as most important. However, if the item does not make sense or is viewed as too simplistic, then the item is rated as less important and will not be ranked. In short, DIT researchers assume that the rating and ranking of items across stories provide an index of the participant's preferred schema and more generally, represent how the participant generally approaches moral decisions beyond the DIT.

How does the DIT measure moral judgment development?

In addition to altering the developmental model underlying the measure, DIT researchers also have focused on how best to define the developmental dimension measured by the DIT. In its original

conception, the DIT assessed a developmental dimension defined in terms of Kohlberg's stages as they were described in the early 70s. More recently, however, the fit of Kohlberg's model to DIT data has been assessed. Based on empirical studies using large and diverse samples including some with as many as 44,000 participants, the description of what the DIT measures have changed.

Specifically, empirical estimates of the ways in which DIT items cluster suggest that the six stages described by Kohlberg do not fit the data. Instead, the obtained number of item clusters suggests three distinct groupings: Stage 2 and 3, Stage 4, and Stage 5 and 6. The finding of three distinct clusters is especially clear when the assessment is based on a heterogeneous sample including participants ranging from high school through the adult years (e.g., Thoma and Rest, 1999). That is, empirically, the best fitting scheme based on DIT data is no longer the six Kohlberg stages. Instead a three level model loosely informed by Kohlberg's model seems more appropriate.

It seems plausible that the obtained clusters are due in part to the adolescent and adult populations typically studied by DIT researchers and perhaps the properties of the DIT itself. However, empirically, it seems clear that participants taking the DIT tend to view items representing Stages 2 and 3 as less important reasoning than items in other clusters. Taken together, the stage 2 and 3 items are not often ranked; although attraction to these items is growing (e.g., Thoma, Bebeau & Dong, in preparation). That is, items that highlight self-preservation, self-interest, and personal relationships are viewed together as personal concerns that are not as central as other more-system wide issues represented by the stage 4 items and those that form the post conventional cluster. Unlike the stage 2 and 3 cluster, the stage 4 and postconventional items are often ranked and viewed as highly important. These findings support the view that the DIT items are assessing moral judgment development at the macro-moral level since the power of the DIT derives from the Stage 4 conventional items and the post-conventional items.

Interpreting the three clusters of items. The three clusters of items suggest that the DIT measures three distinct moral schemas that are developmentally ordered. These schema are labeled: the Personal Interests schema (combining elements of Kohlberg's descriptions of Stages 2 and 3); the Maintaining Norms schema (derived from Kohlberg's definition of Stage 4); and the Post-conventional schema (drawing from Kohlberg's Stages 5 and 6—and equivalent to the items forming the original summary index called the P score). A description of each schema is presented below.

Personal interest schema. Rest, Narvaez, Bebeau, and Thoma (1999) describe the main focus of the personal interest schema as highlighting a perspective that attends the gains and losses each individual may personally experience within a moral dilemma. Similarly, no attention is given to the larger social systems within this schema. Overall, as viewed through a personal interest lens, the social world is a loosely tied network of micro-moral considerations linking close relationships and individual interests. The Personal Interest Schema is fully developed by the time participants are able to reliably complete the DIT (typically defined as a 9th grade reading level). Unfortunately, the DIT can say little about the development of the schema within childhood, except to say that empirically, adolescent and older participants recognized it as, at best, a secondary consideration.

The maintaining norms schema. The Maintaining Norms schema is representative of a society-wide moral perspective. Within the maintaining norms perspective the moral basis of society is understood in terms of how cooperation can be organized on a society-wide basis. However, drawing heavily from the description of Kohlberg's stage 4, the organization of society this schema prioritizes is based on an understanding of rules, roles and the importance of authorities. In addition to Kohlberg's description of stage 4, the Maintain Norms Schema is also informed by Adelson's (1971) conception of the adolescents' developing understanding of political thought and in particular, Adelson's views on adolescent authoritarianism.

More specifically the Maintaining Norms schema has been defined as having the following characteristics: (a) a perceived need for generally accepted social norms to govern a collective; (b) the necessity that the norms apply society-wide, to all people in a society; (c) the need for the norms to be clear, uniform, and categorical (i.e., that there is "the rule of law."); (d) the norms are seen as establishing a reciprocity (each citizen obeys the law, expecting that others will also obey); and (e) the establishment of hierarchical role structures, of chains of command, of authority and duty (e.g., teacher-pupil, parent-child, general-soldier, doctor-patient, etc.—see Rest, Narvaez, Bebeau, & Thoma, 1999, p. 37).

In short, the Maintaining Norms schema prioritizes the established social order and promotes its maintenance as a moral obligation. Consistent with Kohlberg's stage 4, the Maintaining Norms schema support the view that without law there would be no order, people would act on their own special interests with the result a chaotic and lawless society. This schema, does not provide any additional rationale for defining morality beyond simply asserting that an act is prescribed by the law, is the established way of doing things, or is the established Will of God.

Post-conventional schema. Compared to Kohlberg's view of the postconventional stages, DIT researchers assume a different definition of what constitutes a post-conventional system. Avoiding ties to any given philosophical theory or tradition, DIT researchers describe the essential features of Post-conventional thinking in more general terms. In this view, postconventional thinking suggests all moral obligations are to be based on criteria that emphasize shared ideals, are fully reciprocal, and are open to scrutiny (i.e., subject to tests of logical consistency, experience of the community, and coherence with accepted practice--(See Rest, Narvaez, Bebeau and Thoma 1999, p. 38 for a more detailed description).

Based on these descriptions, one can observe that the main source of variance in the DIT is provided by the differences between maintaining norms (conventionality) and Postconventionality. These differences are what Kohlberg regarded as the distinction between Stage 4 and Stage 5; and later Adelson's described as the development of political thought). Although the focus of the DIT measurement system is more directly on the shift from maintaining norms to postconventional thinking than prior models (e.g., Kohlberg's system), the significance of this shift is noteworthy. For instance, the distinction between conventionality and post-conventionality is what tends to drive so many public policy disputes such as the reactions to the wars in Iraq and Afghanistan, how best

to stimulate an economy, minority rights, religion in the schools, medical policy, and so on. Further and perhaps most importantly given the events following 9/11, conventional and post-conventional reasoning addresses the divide between religious fundamentalism and secular modernism (see Marty & Appleby, 1993).

Indices derived from the DIT. For many years, the summary index derived from DIT data was the P score. This score is based on the participant's ranking of post-conventional items. The P score has been criticized for at least two reasons: treating qualitative data as continuous, and for failing to incorporate subject responses to non postconventional items. There is an extensive literature about the first criticism (e.g., Rest, 1979, Rest, 1986, Rest, Thoma, Narvaez, & Bebeau, 1997). In general, the argument advanced by DIT researchers acknowledges the qualitative distinctions between different conceptions of moral thinking as represented by the moral schema. However, the use of a continuous score like the P score, signals the view that the assessment process also is quantitative and should be concerned with the rates of participant responses across the types of moral thinking. P scores, therefore represent the participant's relative location on the developmental continuum (defined by qualitatively different markers). In short, as P scores increase we assume that the participant's developmental location is shifting toward higher levels of moral judgment development.

The second criticism of the P score focuses on the fact that the DIT scoring process does not use all of the participant information available to it. As mentioned previously, P scores only focus on the postconventional schema items and do not attend to other schema scores in the scoring process. The fact that the DIT's main index of development fails to use information on the full complement of schema information has been a concern from many since it violates all of the basic tenants of classical measurement theory (e.g., Loevinger, 1976). Although the P score has been used for many years with general success, there have been a number of attempts to improve on P by supplementing the P score with information from other items. Over the last 10 years, a new index, N2, has been developed and become the primary index of the DIT. The N2 score is best viewed as a modified P score. It uses the P score as its starting point and then adjusts the P score based on the participants' ability to discriminate between P items and lower stage items. The N2 score increases in a positive direction if the individual discriminates high and low items. That is, rates the postconventional items as more important than the personal interest items. Similarly, N2 scores decrease when the participant does not discriminate between postconventional and personal interest items or prefers the personal interest items over the postconventional items. Given that the P and N2 score have a similar starting point, it is not surprising that the correlations between them are high and range from the mid-80s – lower 90s (see Rest, Thoma, Narvaez, & Bebeau, 1997). Empirical comparisons of the two scores indicate that the N2 is an improvement over the P score in older and presumably more developed individuals. Thus, the N2 score should be most helpful in discriminating at the high end of the developmental scale. Current best practice recommendations encourage researchers to use the N2 scores as their summary index when focusing on graduate and professional school populations index as it should be an improvement over the

P scores. By contrast, P scores and N2 scores tend to behave very similarly in high school and college samples suggest that (Office for the Study of Ethical Development, personal communication).

Additional measures derived from the DIT and DIT-2. Beginning in the 1990s there was an interest in developing measures that could broaden our picture of moral judgment development in ways not captured by the schema and summary scores. These measures include an index of developmental phases which describe individuals as either consolidated or transitional in their developmental profile. The second cluster of measures includes indices that can be derived from responses to the DIT and address related but non-moral development constructs. These variables include assessments of social and political attitudes and choices.

Developmental phase indicators. This index was created to explore the role of consolidation and transition on moral judgment development. Thoma and Rest (1999) created a method for assessing developmental phase indicators based upon Snyder and Feldman's (1984) description of developmental phases in development and drawing from Walker and Taylor's (1991) application of the developmental phase notion within the moral judgment domain. Thoma and Rest (1999) measured the degree to which participants were transitional based on a schema profile that indicated little preference for the various stage-based items and, thus, presents a flat response profile. By contrast, a consolidated pattern was indicated when the participant presented a clear preference for a particular schema-based items and, thus, a peaked response profile.

Applications of the developmental phase index were consistent with theoretical expectations. For instance and consistent with the findings reported by Walker & Taylor, (1991), change in moral judgments varied as a function of consolidation and transition Specifically, participants associated with a greater rate of change on DIT summary scores were disproportionately in the group who were moving from a transitional to consolidated phase. Furthermore, Thoma and Rest (1999) found that moral information is more central in the decision-making process during the consolidation phase regardless of developmental level. More recently, developmental phase has been shown to relate to the time it takes to arrive at decisions about moral issues (Thoma, Narvaez, Endicott & Derryberry, 2001). This work found that subjects identified as consolidated took longer to judge the moral issues suggesting a deeper processing of these issues. Further, Derryberry and Thoma (2005) found that developmental phase indicators moderated the link between moral judgment and action. In general, the common finding across these studies is that developmental phase information moderates the relationship between DIT scores and other variables theoretically linked to moral judgment development. These findings indicate that if an effect is observed using the DIT, the same effect will be stronger if computed on participants in the consolidated groupings (Thoma, 2006).

Non-moral judgment measures derived from the DIT. A second set of variables was developed to provide additional information about non-moral constructs by using participant responses to DIT items. For the most part these variables are proxies of non-moral constructs and are useful because they are an efficient way to gather additional information without relying on other measures and the added time demands on participants (Thoma, 2002). At

present these variables capture the following information: a) the degree of decisiveness on the DIT story action choices; b) agreement with action choice decisions made by a group of graduate students in philosophy and political science who achieved the highest scores on the DIT; and c) a proxy measure of religious orthodoxy orientation.

Number of can't decides. The Can't Decide variable is an index of the decisiveness with which an individual selects action choices on the DIT. The procedure used to compute this variable is straightforward and represents a simple count of the can't decide choices. That is, for each of the 6 (or 5 on the DIT2) stories, The DIT asks the participant to choose an action choice for the story protagonist. For example following the Heinz dilemma on the DIT-1, the participant is asked whether Heinz should steal the drug to save his wife or should not steal the drug. A can't decide option is also available. By simply counting the can't decide choices the resulting index ranges from 0 – 6 on the DIT and 0-5 on the DIT-2. The interest in the can't decide index is based on the view that indecision is in part a result of the ease with which participants can process moral information. Following from the Thoma and Rest (1999) study, there is the additional expectation that developmental phase and indecision should be related such that transitional phases should be associated with increased indecision. This expectation is based on the view that transitional phases are associated with multiple and potentially conflicting interpretations of moral situations and issues resulting in more indecision. These expectations have been noted in recent norming studies (e.g., Thoma, Bebeau, Dong, Wiu & Jiang, 2011).

Humanitarian/liberal perspective. The humanitarian/liberal index represents a proxy variable for a humanitarian and liberal perspective on moral issues. This index was created based on the observation that professionals in political science and philosophy obtained the highest P scores for any group similarly assessed. These scores were so high in fact, that for many years this group was used as an "expert" group and used to describe the upper end of the DIT measurement system (Rest, 1979). More recently, and upon closer inspection of this group it was found that not only were these participants obtaining high scores on the DIT, but they were also very consistent in their action choices. As a group these participants supported the position that Heinz should steal the drug for his dying wife. They also endorsed the view that the neighbor should not turn in the escaped prisoner now leading an exemplary life; that the principal should keep the student newspaper open even though they published controversial topics; that the doctor should provide an overdose of a pain killer to a coherent terminally ill patient; that a repair shop owner should hire the minority applicant even if some customers complain and stop patronizing the shop; and that students were justified in occupying the administration building as part of a protest. The clear endorsement patterns suggested a variable in which participants responses to the action choice portion of the DIT assessment is compared to the choices of this "expert" group (Rest, 1979, Thoma 2002). For the DIT -1 the score can range from 0 (no matches) to six (all matches). Across a number of studies the basic finding is that the relationship between moral judgment development and the humanitarian/liberalism scores are curvilinear. The form of

this relationship indicates that high scores are associated with personal interest and post-conventional schemas and lower scores are related to the maintaining norms schema.

Religious orthodoxy. The Religious Orthodoxy score is based on a particular rating and ranking pattern of an item on the doctor's dilemma (or the cancer dilemma on the DIT-2). The particular story containing this item is similar on both versions of the DIT and addresses the question of whether or not the physician ought to provide a drug to a dying woman that will hasten her death. The target item is one that highlights the idea that only God should determine whether one should live or die. By focusing on the ratings and ranking of this item it was noted that a resulting summary index is strongly related to the total scores on religious orthodoxy measures such as the Brown and Lowe Inventory of Religious Beliefs (1951) (Thoma, Bebeau, Dong, Liu, and Jiang, 2011, Narvaez, Getz, Rest, & Thoma,1999).

How do we know the DIT measures moral judgment development?

One historical advantage of the DIT research program is the focus on different approaches to validate a measure of moral judgment development (e.g., Thoma, 2002; Thoma, 2006). Given this focus it is not surprising that the empirical support for the DIT as a measure of moral judgment development are many and varied (see, Rest, 1979, 1986, Rest and Narvaez, 1994, Rest, Narvaez, Bebeau & Thoma, 1999; Thoma, 2006; Thoma, Bebeau, Dong, Liu, & Jaing, 2011). These authors note that a well-articulated set of validity criterion was essential in the development of the DIT. Additionally, these studies contributed to the theoretical shifts mentioned in previous sections. Furthermore, these criterion studies served as the proving ground for new indexes like the N2 score. That is, to support any modification to the scoring or the addition of new indices, the proposed changes were required to yield significantly better trends across criteria and studies than the trends produced by current variables.

These same validity criteria were helpful in addressing criticisms of the DIT. For example, when Sanders, Lubinski and Benbow (1995) concluded that the DIT actually measured verbal ability, DIT researchers were able to find studies that represented the different types of validity criteria and also contained a measure of verbal ability or some reasonable proxy of it (e.g., Thoma, Derryberry, & Narvaez, 2009). The evaluation of these different criticisms was tested by a strategy whereby studies were reanalyze while controlling for verbal ability. Using this approach, the question asked is whether DIT scores can still produce age trends, differentiate known groups, relate to political attitudes and choices and so on when verbal ability is controlled. In response to Sanders, Lubinski and Bebbow's challenge, Thoma, Narvaez, Rest, & Derryberry, (1999) found that when verbal ability was statistically controlled for the dominant trends remained. That is, verbal ability could not account for findings using DIT scores.

The specific criteria used to validate the DIT include: (a) differentiation of various age/education groups; (b) longitudinal gains; (c) correlation with cognitive capacity measures; (d) sensitivity to moral education interventions; (e) correlation with behavior and professional decision making; and (f) predicting to political choice and attitude.

Differentiating age/educational groups. The main approach used in these studies is to assess whether or not the DIT is able to distinguish groups which ought to differ on a measure of moral judgment development. For instance, graduate students in political science and philosophy should score higher than other graduate students who are not so well versed in political and ethical theory. Similarly, college students should score higher than high school students and so on. More recently, large composite samples (thousands of subjects) show that 30% to 50% of the variance of DIT scores is attributable to level of education in samples ranging from junior-high education to Ph.D.s (Thoma, 1986).

The longitudinal gains criteria suggest that a measure of moral judgment development ought to produce evidence of upward movement across time. This criterion follows from the claim that a developmental measure ought to describe change in an upward manner. For instance, a 10-year longitudinal study on the DIT indicates upward change in summary scores for both men and women, for college students and people not attending college, and for people from diverse walks of life (Rest, 1986). A review of a dozen studies comparing freshman to senior college students (n = 755) shows effect sizes (expressed as Cohen's d statistic) of .80 ("large" gains). In short, of all of the variables studied in college student samples, the DIT produces some of the most dramatic longitudinal gains (Maeda, Thoma, Bebeau & You, 2009; Rest & Narvaez, 1994).

Criterion 3 proposes that DIT scores ought to be related to measures of moral comprehension and other cognitive measures. However, relationships with cognitive measures should not be excessive and as such, raise the possibility that DIT scores are actually measuring general cognitive skills. Nor should cognitive measures subsume the relationship between DIT scores and other criterion variables (as claimed by the Sanders, Lubinski, and Benbrow, 1995 study mentioned above). Overall, the existing literature indicates that DIT scores are significantly related to measures of cognitive capacity and moral comprehension, to recall and reconstruction of post-conventional moral argument, to Kohlberg's measure, and to other cognitive developmental measures (Rest, 1979; 1986; Thoma 2006).

The fourth criterion focuses on whether the DIT is sensitive to specific experiences that ought to stimulate development. Intervention studies are the prototype for this criterion (e.g., presence or absence of a dilemma discussion condition). For example, Rest, (1986) describes a review of over 50 intervention studies reports an effect size for dilemma discussion interventions to be .41 ("moderate" gains), whereas the effect size for comparison groups was only .09 ("small" gains).

The fifth criterion suggests that DIT scores ought to be linked to moral actions and desired professional decision making outcomes. For instance, one review reports that 32 out of 47 measures of moral action were statistically significant (Rest, 1986). Furthermore, Rest & Narvaez (1994) linked DIT scores to many aspects of professional decision-making.

Finally, criterion six focuses on the link between DIT scores and social/ political variables. In this cluster, the assumption is that DIT scores should be significantly linked to political attitudes and political choices. This view follows from the position that the DIT is a measure of macro-morality. As mentioned previously, an understanding of macro-morality addresses an understanding of society-wide institutions and their role in promoting social cooperation through laws and the political process. In a review of several dozen correlates between political attitude and DIT scores it was found that they typically correlate in the moderate range (Thoma, Narvaez, Rest, & Derryberry, 1999, Crowson, DeBacker, & Thoma, 2005). When DIT scores were combined in multiple regression with measures of cultural ideology, the overall prediction increased to up to two-thirds of the variance in opinions about controversial public policy issues. These issues include abortion, religion in the public school, women's roles, rights of the accused, rights of homosexuals, civil liberties, the rights of minorities, and free speech issues. Given that these issues are among the most hotly debated of our time, the DIT has the potential to contribute to our understanding of individual differences in political preferences and attitudes.

In addition to these validity criteria, DIT researchers also focused on traditional standards for tests and measures such as acceptable psychometric evidence as well as response stability across different test-taking sets. In addition, DIT scores show discriminate validity from a host of competing variables such as verbal ability/general intelligence and from conservative/liberal political attitudes (Thoma, Derryberry, & Narvaez, 2009; Thoma, Narvaez, Rest & Derryberry, 1999). Moreover, the DIT is equally valid for males and females since gender accounts for less than one half of a percent of the variance of the DIT, whereas education is 250 times more powerful in predicting DIT variance (Thoma, 1986).

» SUMMARY

The DIT has evolved significantly over its 35-year history and from its roots in the Kohlbergian model. Consistent with other contemporary models of development the shift to a neo-Kohlbergian position resulted in some modifications and rejections of traditional assumptions. Unlike other models that expanded the measurement system to broaden their focus beyond moral functioning (e.g., the MHC approaches), DIT researchers maintained their interest in moral functioning. Specifically, the DIT is claimed to measure default schema by which individuals interpret moral issues. Focusing on the macro-moral level, these default schemas inform the individual's understanding of social structures and their mechanisms. Further, it is claimed that the development of these schemas is ordered such that, starting during the second decade of life, a focus on understanding and maintaining norms gives way to a post-conventional understanding.

These assumptions have been supported by the six validity criteria clusters that contain multiple indicators and cohorts. The results of these analyses clearly support the view that the DIT measures a developmental construct within the moral domain. Further, existing evidence suggests that the measure is particularly good at assessing the shift from a conventional/maintaining norms perspective to a post conventional view of social cooperation. These findings suggest that the DIT will continue to offer the field a theoretical model and research strategy that serves to further moral judgment research. ■

REFERENCES

Adelson, J. (1971) The political imagination of the young adolescent. *Daedalus*, *100*(4), 1013-1050.

Brown, D. G., & Lowe, W. L. (1951). Religious beliefs and personality characteristics of college students. *Journal of Social Psychology, 33*, 103-129. doi:10.1080/00224545.1951.9921803

Commons, M.L. & Pekker. (2007). Hierarchical *Complexity: A formal Theory*. Unpublished manuscript submitted for publication.

Crowson, H. M., DeBacker, T., & Thoma, S.J. (2007). Are DIT scores empirically distinct from measures of political identification and intellectual ability? A test using post-9/11 data. *British Journal of Developmental Psychology, 25*(2), 197-211.

Derryberry, P. & Thoma, S.J. (2005). Moral Judgment, Self-Understanding, and Moral Actions: The Role of Multiple Constructs. *Merrill-Palmer Quarterly, 51*(1)*,* 67-92.

Kohlberg, L. (1969). Stage and sequence. The cognitive- developmental approach to socialization. In Goslin (ed.) *Handbook of socialization theory and research*. Chicago: Rand McNally.

Kohlberg, L. (1984*). Essays on Moral Development: The nature and validity of moral stages, Vol. 2*. San Francisco: Harper & Row.

Loevinger, J. (1976). *Ego Development*. San Francisco: Jossey-Bass Publishers.

Maeda, Y., Thoma, S. J., & Bebeau, M. (2009). Understanding the relationship between moral judgment development and individual characteristics: The role of educational contexts. *Journal of Educational Psychology, 101*(1), 233-247. doi:10.1037/a0013381

Marty, M. E., & Appleby, R. S. (Eds.), (1993). *Fundamentalism and the State*. Chicago: University of Chicago.

Narvaez, D. & Bock (2002). Moral Schemas and Tacit Judgments or How the Defining Issues Test is Supported by Cognitive Science. *Journal of Moral Education*. *31*(3), 297-314. doi:10.1080/0305724022000008124

Narvaez, D., Getz, I., Rest, J., & Thoma, S. (1999). Individual Moral Judgment and Cultural Ideologies. *Developmental Psychology*, *35*(2), 478-488.

Rest, J., Cooper, D., Coder, R., Masanz, J., & Anderson, D. (1974). Judging the important issues in moral dilemmas — an objective test of development. *Developmental Psychology, 10*(4), 491-501. doi:10.1037/h0036598

Rest, J. (1979). *Development in judging moral issues*. Minneapolis, MN: University of Minnesota Press.

Rest, J. (1983). *Morality*. In P. H. Mussen (Series Ed.) & J. Flavell & E. Markman (Vol. Ed.), *Handbook of child psychology: Vol. 3, Cognitive Development* (4th ed., pp. 556-629). New York: Wiley.

Rest, J. (1986). *Moral development: Advances in research and theory.* New York: Prager.

Rest, J. & Narvaez, D. (1994). *Moral development in the professions*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rest, J., Narvaez, D., Bebeau, M., & Thoma, S. (1999). *Post-conventional moral thinking: A Neo-Kohlbergian approach*. Mahwah, NJ: Lawrence Erlbaum Associates.

Rest, J., Thoma, S. J., Narvaez, D., & Bebeau, M. J. (1997). Alchemy and beyond: Indexing the Defining Issues Test. *Journal of Educational Psychology, 89* (3), 498-507.

Rummelhart, D.E. (1980) Schemata: the building blocks of cognition, In R. Spiro, B. Bruce, & W. Brewer (Eds.), *Theoretical Issues in Reading Comprehension*, (pp. 33-58), Hillsdale, NJ: Erlbaum.

Sanders, C.E., Lubinski, D., & Benbow, C.P. (1995). Does the Defining Issues Test measure psychological phenomena distinct from verbal ability? *Journal of Personality and Social Psychology, 69*(3), 498-504.

Snyder, S. S., & Feldman, D. H. (1984). Phases of transition in cognitive development: Evidence from the domain of spatial representation. *Child Development, 55*(3), 981-989. doi:10.1111/1467-8624.ep12427036

Thoma, S. J. (1986). Estimating gender differences in the comprehension and preference of moral issues. *Developmental Review, 6*(2)*,* 165-180. doi:10.1016/0273-2297(86)90010-9

Thoma, S. J. (2006). Research using the Defining Issues Test. In Killen and Smetana (Eds.), *Handbook of Moral Psychology.* L. Earlbaum: Mawah, NJ.

Thoma, S. J. (2002). An overview of the Minnesota approach in moral development. *Journal of Moral Education*, *31*(3), 225-246. doi:10.1080/0305724022000008098

Thoma, S. J., Derryberry, P. & Narvaez, D. (2009). The distinction between moral judgment development and verbal ability: Some relevant data using socio-political outcome variables. *High Ability Studies*, *20*(2), 173-186. doi:10.1080/13598130903358535

Thoma, S. J., Narvaez, D., Rest, J. & Derryberry, P. (1999). The distinctiveness of moral judgment. *Educational Psychology Review, 11*, 361-376.

Thoma, S. J., Bebeau, M. J., & Dong, Y. Moral Judgment competency is declining over time: Evidence from 20 years of Defining Issues Test data. In preparation.

Thoma, S. J, Narvaez, D., Endicott, L. & Derryberry, P. (2001). *Developmental phase indicators and moral information processing.* Paper presented to the American Educational Research Association, New Orleans, LA.

Thoma, S.J., Bebeau, M. J., Dong, Y., Liu, W., & Jiang, H. (2011). Guide for DIT-2. Available from the Office for the Study of Ethical Development.

Thoma, S. J., & Rest, J. (1999). The relationship between moral decision-making and patterns of consolidation and transition in moral judgment development. *Developmental Psychology, 35*(2)*,* 323-333. doi:10.1037/0012-1649.35.2.323

Torney-Purta, J. (1990). *Youth in relation to social institutions*. In S. Feldman & G. R. Elliott (Eds.), *At the threshold: The developing adolescent* (pp. 457-478). Cambridge, MA: Harvard University Press.

Walker L. J. & Taylor, J. H. (1991). Stage Transitions in moral reasoning: A longitudinal study of developmental processes. *Developmental Psychology*, *27*(2), 330-337. doi:10.1037/0012-1649.27.2.330

Hierarchical complexity in physics

Kristian Stålne[1], Michael Lamport Commons[2], and Eva Yujia Li[3]

1 Lund University
2 Harvard Medical School
3 Harvard Graduate School of Education

ABSTRACT

The derivation of string theory from the two paradigms of wave theory and of relativity is a stage 14 task. The wave theory may partially be represented by the acoustic wave equation for a fluid in one dimension. The stages of development of the wave equation can be is presented in terms of increasing orders of hierarchical complexity. The derivation, shown from order 9 concrete to 14 paradigmatic is presented as a schema where it is specified how a higher order is created by coordinating elements from the respective previous order. The wave equation at the paradigmatic order is created by coordinating the three metasystematic relationships: Newton's Law of Motion, the Constitutive equation and the Ideal gas law. These three relationships in turn coordinate the variables force, density and acceleration, all being systematic since they are functions of time and location. This result gives an understanding of how knowledge is organized in the acoustic domain and in adjacent domains such as classical and solid mechanics. This paradigm is also combined with notions from general relativity to show that the two paradigms may be combined to form a crossparadigmatic task. One result is string theory. It also serves as an illustrative example of the principles of MHC.

KEYWORDS: string theory, wave theory, relativity, orders of hierarchical complexity, crossparadigmatic task

The purpose of this paper is to apply the Model of Hierarchical Complexity (Commons, 2008; Commons, Pekker, 2008; Commons, Trudeau, Stein, Richards, & Krause, 1998) to determine the stage of physics theories. It is an analysis using the Model of Hierarchical Complexity (MHC) of the historical development of String theory and its alternatives and antecedents starting fluid mechanics and their wave equation. The MHC explains why the tasks were ordered as they were and had to unfold in the order they did. But the time, place and person of the development of each advance was not constrained by the MHC and is chaotic. It will show how the wave equation for a fluid in one dimension is derived by increasing orders of hierarchical complexity. It will also be illustrated how the string theory is an Order of Hierarchical Complexity 15 task as a coordination of the laws of quantum mechanics and the theory of relativity. It is important to clarify that the tasks for which the orders of hierarchical complexity will be determined are the tasks of formulating theories. The paper takes perspective of the scientists who formulated the theories. It takes into account what was known at the time the theories were created, and what knowledge the scientists built on to formulate their theories. These facts are revealed by the history of science. This is in contrast with the task of understanding or applying the theories. Once the theories have been formulated, the task of understanding and applying them drops down in stage because of what has been termed support (Fischer, et al, 1984).

The paper first introduces the Model of Hierarchical Complexity and the orders that are derived from it. Then it is demonstrated how the wave equation is being built up through the increasing orders of complexity. The resulting wave equation and its generalization to quantum physics are found to be of the 14th paradigmatic order. String theory, which reconciles quantum physics and the theory of general relativity, is at the 15th cross-paradigmatic order. In order to move up to the 16th order, one has to be able to reflect on actions on the 15th order.

» THE MODEL OF HIERARCHICAL COMPLEXITY

The Model of Hierarchical Complexity is a measurement system that sets forth the hierarchical structure by which actions are put into a hierarchical order. Actions are behavioral events that produce outcomes. A task is a set of required actions to obtain an objective. In the literature, two types of complexities have been identified (Commons, Trudeau, et al, 1998): horizontal (traditional) and vertical (hierarchical). In traditional horizontal complexity, the complexity of a task is determined by the number of times a specific subaction is repeated. In hierarchical complexity, the complexity of an action is determined by the non-arbitrary way in which the subactions are organized, not how many subactions there are (Commons & Pekker, 2008).

Author note: Correspondence regarding this paper should be addressed to Dr. Kristian Stalne, Department of Construction sciences, Lund University. E-mail: kristian.stalne@construction.lth.se

In hierarchical complexity, actions at a *higher order of hierarchical complexity* (see figure 1)*:*

a) Are defined in terms of actions at the *next lower* order of hierarchical complexity

b) Organize and *transform* the lower-order actions

c) 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, we say the higher-order action *co-ordinates* the actions of the next lower order. The task of evaluating $a \times (b + c)$ is used as an example. The standard way to complete this task is to distribute *a* over *b* and *c* is by having $(a \times b) + (a \times c)$. This shows how distribution is built out of the actions of $\times$ and $+$. Contrast this to the case of $(a + b) + c$. Addition is associative and $(a + b) + c$ is equivalent to $(a + b) + c$ or $a + (b + c)$. Therefore, in the task of $(a + b) + c$, the organization of two actions of addition is arbitrary. So evaluating $a \times (b + c)$ is more hierarchically complex than the task of evaluating $(a + b) + c$. The task of distributing is also more hierarchically complex than the two-part task of first evaluating $b + c = d$ and then evaluating $a \times d$.

Actions of each stage coordinate actions that are one stage lower, thus creating a hierarchical system. *Stage of performance* is defined as the highest-order hierarchical complexity of the task solved (Commons, Miller, Goodheart, & Danaher-Gilpin, 2005).

In previous research, tasks have been found to occur at 17 orders of hierarchical complexity, from 0 (calculatory) to 16 (metacross-paradigmatic). Table.1 shows the orders of Hierarchical Complexity. Thus far, however, there have been few examples of tasks at the 15 cross-paradigmatic order. The order sequence presumably is infinite, but because of human limitations, we have created only 16 (and possibly 17) orders.

The wave equation for a fluid in one dimension

The wave equation describes the behavior of waves in a medium. The following equation is the wave equation of a pressure wave in a fluid in one dimension, where *p* is pressure and *c* is the speed of sound in the fluid.

$$\frac{\partial^2 p}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0 \tag{1}$$

In the following it will be demonstrated how the wave equation in one dimensional is derived by coordinating more and more complex building blocks, from the 9 concrete order, through each of the following orders, arriving at the final result, the wave equation at the 14 paradigmatic order.

» ORDER 9 CONCRETE

At the concrete order a particle's or fluid element's state can be given in terms of actual numbers that represent

» Displacement $u_1, u_2, u_3, \ldots$
» Particle velocity $v_1, v_2, v_3, \ldots$
» Particle acceleration $a_1, a_2, a_3, \ldots$
» Pressure $p_1, p_2, p_3, \ldots$
» Density $\rho_1, \rho_2, \rho_3, \ldots$

The respective state can be given at different certain times $t_1, t_2, t_3, \ldots$ and at a certain positions $x_1, x_2, x_3, \ldots$. The subscripts indicate that these variable values are actually specific instances. In logic, they are called specified variable and therefore are concrete.

» ORDER 10 ABSTRACT

Abstract variables are created by coordinating every possible outcome of the concrete instances, specific times or specific positions. The abstract state variables, or field variables as they are referred to, are used that coordinates every possible displacement *u*, velocity *v*, acceleration *a*, pressure *p* and density ρ (rho). At the abstract order, time and position are expressed as variables *t* and *x*.

At the abstract order, the equation for pressure is provided. Even though the equation is from the systematic order, the variables themselves can be viewed as just variables. The equation is given and all a participant has to do is to put in the correct values for the derivatives. The definition provided for change force *dF* per change in unit area *dS* is

$$p = \frac{dF}{dS} \tag{2}$$

The definition provided for density ρ (rho) is change in mass *dm* per change in unit volume *dV*

$$\rho = \frac{dm}{dV} \tag{3}$$

» ORDER 11 FORMAL

At the formal order, field variables are expressed as single variable functions or deduced by derivation with respect to one variable. The very notion of a function, a mapping relating two variables, input and output, is always formal.

At a fixed location $x = x_0$,

$$\begin{aligned}
u = u\,(t) &= \hat{u}\cos(\omega t) \\
v = v\,(t) &= \hat{v}\cos(\omega t) \\
a = a\,(t) &= \hat{a}\cos(\omega t) \\
p = p\,(t) &= \hat{p}\cos(\omega t) \\
\rho = \rho\,(t) &= \hat{\rho}\cos(\omega t)
\end{aligned} \tag{4}$$

A letter such as *u* with the symbol is read "u hat". $\hat{u}, \hat{v}, \hat{a}, \hat{p},$ and $\hat{\rho}$ are the constant amplitudes and ω (omega) is the angular velocity which relates to the period time T according to $\omega = 2\pi/T$.

Table 1. Orders of hierarchical complexity

Order	Name complexity
0	Calculatory
1	Automatic
2	Sensory & 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	Cross-paradigmatic
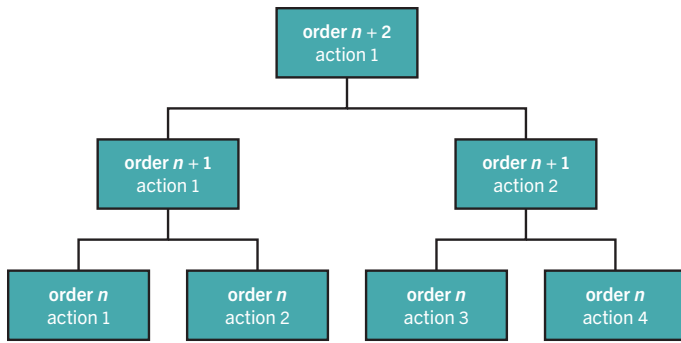
Figure 1. Hierarchical Structure of Tasks

At a fixed time $t = t_o$ the pressure in one dimension is

$$p = p(x) = \hat{p}\cos(kx) \tag{5}$$

This corresponds to a snapshot of the pressure of a one dimensional travelling wave in a duct, such as an air shaft, as a function of the position x. k is the constant wave number, which can be seen as a spatial angular frequency with is related to the wavelength λ (lambda) according to $k = 2\pi/\lambda$.

» ORDER 12 SYSTEMATIC

At the systematic order 11, the field variables are expressed as functions of more than one abstract variable. Here the field variables u, p, ρ, etc are expressed as functions of both time and location, according to,

$$\begin{aligned} u(x,t) &= \hat{u}\sin(kx - \omega t) \\ p(x,t) &= \hat{p}\cos(kx - \omega t) \\ \rho(x,t) &= \hat{\rho}\sin(kx - \omega t) \end{aligned} \tag{6}$$

Kinematics describes the movement of particles expressed in particle displacement, velocity and acceleration. Velocity and acceleration are defined as the derivatives of displacement and velocity, respectively, with respect to time according to

$$\begin{aligned} v(x,t) &= \frac{\partial}{\partial t}(u(x,t)) \\ a(x,t) &= \frac{\partial}{\partial t}(v(x,t)) \end{aligned} \tag{7}$$

The field variables, which are functions at the systematic order, can be grouped into the three categories of kinematics - expressed in displacement, velocity or acceleration, force - expressed in pressure, and mass - expressed in density.

» ORDER 13 METASYSTEMATIC

The metasystematic order 13 is characterized by coordination of two or more systems at the systematic order.

Kinetics is achieved by means of Newton's law of motion in rigid body dynamics, which is the coordination of kinematics and force. The derivation of Newton's law of motion for a fluid is therefore a coordination at the Metasystematic order, since it successfully coordinates the system of force through the pressure $p = p(x, t)$

with the system of kinematics through the acceleration $a = a(x, t)$. Using pressure $p(x, t)$ and particle velocity $v(x, t)$, Newton's law of motion for a fluid in one dimension can be expressed as,

$$\rho_0 \frac{\partial v}{\partial t} = -\frac{\partial p}{\partial x} \tag{8}$$

where ρ_o is the mean density of the fluid.

Another example of a metasystematic coordination is the *Continuity Equation*, which is based on the principle of indestructibility of mass. It is a mathematical formulation of the relationship between changes in density $\rho(x, t)$ and changes in volume of an element, which can be expressed with the particle velocity of the element $v(x, t)$, as a function of time and position according to

$$\frac{\partial \rho}{\partial t} = -\rho_0 \frac{\partial v}{\partial x} \tag{9}$$

A third example of a metasystematic coordination is the *Ideal Gas Law*, which gives a relationship between the pressure $p(x,t)$ and the density $\rho(x, t)$. From the ideal gas law the following equation can be derived, where the right hand side only contains constants.

$$\frac{\partial p}{\partial \rho} = \kappa \frac{p_0}{\rho_0} \tag{10}$$

It can be noted that these three examples of relationships at a metasystematic level coordinates the variables that reflects different aspects of the phenomenon, or categories, a wave motion studied as a propagation of force, displacement and mass.

» ORDER 14 PARADIGMATIC

At the Paradigmatic order, the wave equation is derived by coordinating the three metasystematic relations presented above:

» Newton's law of motion
» The Continuity Equation (Conservation of Mass)
» The Ideal Gas Law

The coordination is performed by employing the three Metasystematic relationships to eliminate two of the field variables, usually velocity and density, to achieve the final result, the wave equation expressed in pressure $p(x, t)$ as a field variable,

$$\frac{\partial^2 p}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0 \tag{11}$$

The solution that satisfies the wave equation describes a travelling wave propagating in a fluid in the positive (first half) and the negative (second half) x-direction.

$$p(x,t) = \hat{p}_+ \cos(kx - \omega t) + \hat{p}_- \cos(kx + \omega t) \tag{12}$$

The same type of relation can be derived for waves in solid media, where shear forces and torques also have to be considered. This will result in not only longitudinal waves but shear, bending, rotational and surface waves as well.

The basic derivation and appearance of the wave equation will be similar for all cases. For example, the classical wave equation, which models a wave on a disturbed string, is expressed by the following, where y is the vertical displacement of a given point at the position x on the string.

$$\frac{\partial^2 y}{\partial t^2} - v^2 \frac{\partial^2 y}{\partial x^2} = 0 \qquad (13)$$

All the wave equations are at the 14th paradigmatic order.

The wave equation also shows up in quantum mechanics. The time-independent Schrödinger Equation is a decoordination of classical wave equation and the conservation of energy, Total Energy. $TE = PE + KE$, where PE = potential energy, KE = kinetic energy. The coordination is at the paradigmatic stage because the conservation of energy is at the Metasystematic stage. The coordination of an action at the paradigmatic stage and an action of metasystematic stage completes a task at the paradigmatic stage. The following equation is a one-dimensional, time-independent Schrödinger Equation for a particle of mass m, commonly known as the Time-Independent Schrödinger Equation,

$$-\frac{h^2}{2m} - \frac{d^2 y}{dx^2} + V(x)\, y = Ey \qquad (14)$$

where m is the particle mass, y is the vertical displacement of a particle, and $V(x)$ is the potential energy of a particle as a function of position x.

» ORDER 15 CROSS-PARADIGMATIC

At the cross-paradigmatic order, the field of quantum mechanics is reconciled with the theory of general relativity. To understand the reason that this coordination is at the 15th order, it is helpful to review its history and background of the theory of relativity.

Special and general relativity

Albert Einstein (1950) created a new model of the universe by coordinating the paradigm of the theoretical and experimental result that light travels in a constant speed with the paradigms of classical physics to form the field of relativity.

In the field of electromagnetism, Maxwell's equation gives the result that the speed of light has to be the same to all the time (Toth, 2003). This result conflicts with the laws of classical mechanics. According to Newton's classical mechanics, the speed of a moving object is observed to be different by observers moving at the different speeds. They observe the relative speed of the moving object compared with themselves. The constant speed of light in that theory is paradoxical, because it seems to suggest that the speed of observer does not matter. Maxwell explained this by proposing another theory. He proposed that light has to be transmitted by a type of medium, which he named "ether", that the universe is full of. Ether is static in the universe. As the earth revolves around the sun, it moves crossing the "ether field". Maxwell proposed that the speed of light solved by Maxwell's equation is the "absolute" speed of light in the universe. However, as the earth revolves around

the sun, there should exist "relative" speed of light. According to this proposed theory, the speed of light is relative to the speed of the observer.

Michelson and Morley (1887) tested the existence of ether by measuring the speed of light at a static point and at a moving point. Surprisingly, this experiment showed that the speed of light is the same whether or not the observer is moving, disconfirming Maxwell's theory of ether.

Einstein realized that to accept the speed of light as being constant regardless of the position and speed of the observer is to establish a new space -time model of the universe. He derived the theory of special relativity by keeping the speed of light constant and making time and space flexible. His theory suggested that time and space are contractible. An observer on a fast moving spaceship experiences time slower and space shorter than the observer on a slowly moving spaceship. This theory has been confirmed by experiments, such as showing that the amount of energy goes up as a particle is accelerated towards the speed of light. Einstein created a four-dimensional framework of the universe, three dimensions of space and time. Later, the theory of special relativity was expanded to the theory of general relativity. It made it possible for Einstein to explain gravity and its equivalence to momentum. It also predicted that light would appear to be bent when it passed near the Sun. This is because space time is warped or curved by the mass of the Sun.

The four dimensional space equations are described below. Defining the event to have space-time coordinates (t, x, y, z) in system s and (t', x', y', z') in s', then these coordinates are related in the following way:

$$\begin{aligned} t' &= \gamma\left(t - \frac{vx}{c^2}\right) \\ x' &= \gamma(x - vt) \\ y' &= y \\ z' &= z \end{aligned} \qquad (15)$$

where

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad (16)$$

is the Lorentz factor, c is the speed of light in a vacuum and v is the speed of the system S' relative to S.

The reasons that the coordination is at the cross paradigmatic order 15 is that the relativity theories coordinated two paradigms at order 14. The first paradigm is the model of light waves propagating at a constant speed in vacuum. This is the theoretical conclusion of Maxwell's equations and the experimental result of Michelson and Morley. Understanding the theoretical implication of Maxwell's equation is a task at order 14. Maxwell's equations coordinated two order 13 metasystems, electric field and the magnetic field. Understanding the empirical evidence provided by the experiment was correct is a task at order 12. Einstein abandoned the theory of ether, which marked the end of the early 19th century paradigm of physics in which all waves had to travel in a medium.

The second paradigm is the old paradigm of Newton's laws of mechanics positing a gravitational field within Euclidean geometry and founding mathematical physics. It has distance independent of time and of rate. Understanding the interrelatedness of all of Newton's laws and the properties of the system in classical physics is a task at order 13. These two paradigms had intrinsic conflict with each other concerning the speed of light. Einstein reconciled the two by constructing a new cross-paradigmatic theory in which time, distance, an even mass, are all transformed by showing that all of them are a function of their relative speed with respect to the speed of light. The extension to the general relativity theory, integrates space-time with inertia and gravity.

Because mass in their as in $E = mc^2$, this integrates a new physics, geometry.

String theory

Quantum mechanics describes the properties of particles at the subatomic level. It describes the subatomic particles as operating with uncertainty and probability, according to the Heisenberg uncertainty principle. Quantum mechanics successfully explains three of the four fundamental forces in physics, the strong force, the weak force and the electromagnetic force. According to quantum mechanics, forces are created by the exchange of messenger particles. For example, electromagnetic force is created by the exchange of photons. The more exchange, the stronger the forces. However, this theory does not explain the last fundamental force of physics, gravity (Zwiebach, 2004).

Einstein's theory of general relativity describes the gravity as a function of space and time. The theory of general relativity is an improvement over Newton's immutable mechanics. It describes the universe as mechanical and predictable. This theory can be observed with massive objects, but not with the microscopic particles, characterized by chaotic movements and unpredictability.

Both theories have been experimentally tested and proven valid. However, general relativity and quantum mechanics seem to be incompatible with each other. They paint distinct pictures of the universe – one operates under mechanical laws and the other filled with uncertainty. They also have disjointed experimental domains. General relativity is only observable with massive objects. Quantum effects are only observable with minute particles. Could there be a single unified theory that explains the universe on both the macro and on the micro scale?

No theory to date has successfully reconciled quantum mechanics and general relativity. However, there are a few plausible working models. String theory is an active research framework in the field of physics (Polchinski, 1998). It proposes that everything in the universe is composed of tiny vibrating strings. The shape of the string and the way that the strings vibrate contribute to matters' unique properties, such as mass. The string theory describes that gravity is produced by one type of vibrating string called the graviton. It offers an explanation of how gravity works in the subatomic scale. This is the key to unifying the four forces, gravity, the strong force, the weak force, and electromagnetic force.

There are also other alternatives to string theory that unite the two camps. Examples are Loop Quantum Gravity and Quantum Gravity.

These theories are at the cross-paradigmatic order 15 because they successfully coordinate two theories at the paradigmatic order.

» ORDER 16

At order 16, the action required in the transition is to reflect on Order 15 tasks. Scoring order 15 tasks is, but not completely, an order 16 task. It is in transition because one has to be at a higher stage in order to score the lower stages. An order 16 tasks requires a reflection on a stage 15 task and what is missing from it. The reason that it is transitional is that there is not a positive description of the order and how it coordinates two or more cross-paradigmatic order tasks. The order sequence presumably is infinite, but because of human limitations, we have created only 15 and possibly 16 Orders.

» CONCLUSION

The derivation of the wave equation for a fluid in one dimension serves as an illustrative example of the principles of the Model of Hierarchical Complexity. Lower stage elements are coordinated by higher stage systems, and the systems become increasingly more complex hierarchically. It is shown that the resultant wave equation is at the paradigmatic order 14. The wave equation is generalized to describe wave properties of particles in the quantum realm. At the cross-paradigmatic order 15, quantum physics and the theory of general relativity are reconciled by the string theory. At the next order above, the ability to reflect on order 15 task is attained. This result gives an understanding of how knowledge at the highest known orders of human performance is organized. This result can be generalized to other domains and support progress in areas that have not yet reached that high in complexity. ∎

REFERENCES

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., Bresette, L. M., & Ross, S. N. (2008). The connection between postformal thought and major scientific innovations. *World Futures: Journal of General Evolution, 65*(1-3), 503-512.

Commons, M. L., & Pekker, A. (2008). Presenting the formal theory of hierarchical complexity. *World Futures: Journal of General Evolution, 65*(1-3), 375-382.

Commons, M. L., Trudeau, E. J., Stein, S. A., Richards, F. A., & Krause, S. R. (1998). The existence of developmental stages as shown by the hierarchical complexity of tasks. *Developmental Review, 8*(3), 237-278.

Einstein, A. (1905). On a heuristic viewpoint concerning the production and transformation of light. *Annalen der Physik 17*(6): 132–148. doi:10.1002/andp.19053220607

Michelson, A. & Morley, E. (1887), "On the Relative Motion of the Earth and the Luminiferous Ether". *American Journal of Science, 34*, 333–345.

Polchinski, J. G. (1998). String theory. Cambridge university press.

Skudrzyk, E. (1971). *The Foundations of Acoustics – Basic Mathematics and Basic Acoustics.* Verlag, Wien, Austria: Springer

Toth, V. (2003). Why is the speed of light constant? Retrieved from http://www.vttoth.com/LIGHT/light.htm

Zwiebach, B. (2004). *A first course in String Theory.* Cambridge university press, New York.

BEHAVIORAL DEVELOPMENT BULLETIN

Review and shortcomings of literature on corruption in organizations in offering a multi-faceted and integrative understanding of the phenomenon

Elke Fein and Jürgen Weibler

University of Hagen, Germany

ABSTRACT

This article provides a brief overview of literature on corruption from different disciplinary perspectives. After a short look at contributions from history, sociology, anthropology and psychology, the paper primarily reviews articles on corruption in organizations from fields like organizational behavior (OB), behavioral ethics (BE) and management studies (MS). Despite frequent calls for a more interdisciplinary or even a "holistic view" of corruption in this literature, we claim that the literature reviewed here often fails to offer an adequate, i.e. multi-faceted and integrative understanding of the phenomenon, and that this is due to disciplinary constraints and traditions often inducing researchers to take less-than-desirably complex views onto the phenomenon. Moreover, we argue that many articles on corruption do not reflect, question and/or contextualize their own moral and/or ethical standards and evaluation criteria systematically. This is shown, first, with regard to the degree of reflexivity of the applied analytical terms and concepts in general and with regard to the extent to which value judgments are contextualized in particular. Second, our claim is illustrated by a tendency to underrate or ignore major aspects of the subjective dimension of behavior, namely actors' empirical action logics.

KEYWORDS: corruption, contextualization, organizational behavior, management studies, behavioral ethics

It has been repeatedly acknowledged that scholarly interest in ethical issues has grown in recent years. It is therefore not surprising that corruption has become a focus of study in many social science disciplines each of which, we claim, has important contributions to make. In view of proposing an integrative, interdisciplinary framework for understanding and explaining corruption, as well as attitudes towards corruption (see Fein & Weibler, 2014), the following paper provides a non-exhaustive overview of literature on corruption from different disciplinary perspectives with a special focus on their respective structural complexity and self-reflexivity. It begins by briefly referring to some of the most important "classic" social science perspectives such as history, sociology, and anthropology, giving a short summary of their central outlooks on and findings about corruption,

as well as of the insights to be gained from them in view of a more systematic, integrative account of corruption and unethical behavior. The second, more detailed section of this paper focuses on how corruption is dealt with by different strands of organization studies, amongst others within Behavioral Ethics, Organizational Behavior and Management Studies (BE/OB/MS).

Note that our review of either of the fields considered cannot give encompassing or representative overviews of the research on corruption done in the respective fields. We would therefore like to stress that the main intention of this paper is not to give comprehensive evaluations, but rather to identify some of the central, typical features of looking at the problem in each of the disciplines considered, with a particular focus on the scope of their typical perspectives and the structure of their most frequently used patterns of argumentation. On the whole, we are looking for valuable insights, as well as for potential shortcomings, limitations and reductionisms which might be overcome by more integrative perspectives on corruption and thus, a more complex and more effective corruption analysis, research, and practice.

» CORRUPTION VIEWED THROUGH THE LENSES OF HISTORY SOCIOLOGY, ANTHROPOLOGY AND PSYCHOLOGY

Before turning to research on behavioral ethics in organizations, we take a brief look at what we think are the most valuable thoughts and contributions other social science disciplines have to make to a more integrative theoretical endeavor. The following paragraphs in this section are primarily intended to extract those contributions.

Historians are interested in historical phenomena and the dynamics of their development. They have thus not only described how corruption itself has historically changed, but have also started to study the emergence and the changes of value systems defining what was considered to be a legitimate and/or, in contrast, corrupt action across long periods of time (Engels, Fahrmeir & Nützenadel 2009). As one of the first researchers, Joseph A. Senturia stated that the understanding of the term corruption depended on the opinion of the respective observer and on the dominant political and public morality (Senturia, 1930). Michael Johnston therefore suggests "that we use the concept of corruption to ask questions about state, society, and political change" rather than about particular behavior. For "corruption is a political and normative concept rather than a kind of 'natural' category of unacceptable action" (Johnston, 2005, p. 71-72). Many historians have come to understand what Vadim Volkov (2000) has called the "historical relativity of corruption" and its connectedness to "a specific type of social organization, the state". They therefore mostly interpret corruption as a typical product of modernization. For example, Jens Ivo Engels, one of the leading German scholars on historical corruption, claims that the classic definition of corruption as a misuse of public office for private gain "only makes sense within modern societies" (Engels, 2010), while in pre-modern societies, where public and private spheres had not yet been differentiated, it was common and thus normal to hold and treat offices as a means of personal enrichment. Engels therefore urges to distinguish, first, between the modern scientific notion of corruption and that of the respective times. Second, he suggests to distinguish between practices of and debates about corruption. While pre-modern times knew neither modern morality nor the differentiation of spheres necessary to engage in anti-corruption discourses and/or practices, the latter almost automatically contain a moral judgment typical of modern, self-reflexive discourses (Engels, 2010).

Similarly, Werner Plumpe (2009), referring to James Cameron Scott's classic *Comparative Political Corruption* (1972) claims that "corruption and modernity are co-evolving" phenomena, since only modernity has set up extensive judicial rules governing economic life, while common behavior was not regulated before (no rule, no crime). In fact, Engels explains the scandalization and criminalization of corruption as opposed to civilization as a result of the intellectual quest for clear evaluations and categorizations which he sees as a typical feature of modern ambitions to "clean" public thinking, as well as social life, from ambiguities.

However, historical accounts have also observed that large parts of the population in most countries were not prepared to meet the moral demands of modernization (Engels, 2010) – a finding which still holds true for contemporary transforming societies and social organizations and will be further discussed in relation to psychological, developmental theories below (see Volkov 2000).

Likewise, Sociologists typically interested in the emergence and acceptance of social norms and in individual behavior which either conforms to or diverges from those norms, similarly claim that corruption has social and cultural roots. Like historians, sociological perspectives therefore stress that "corrupt" behavior is not always considered as being unethical and divergent, but it rather still constitutes the norm in many social contexts today (Fleck & Kuzmics, 1985). Many sociologists who look at corruption through the lens of sociological scholarship have, for example, observed the same patterns of behavior (nowadays commonly evaluated as "corrupt") in developing, such as, modernizing third world countries as in historical pre-modern societies, in the context of the Italian mafia (Arlacchi, 1989) or in socialist systems such as the Soviet Union (Voslensky,1987). With regard to 17th/18th century England historical corruption research speaks of "protocorruption" (Scott 1985), exactly because the respective phenomena were not considered problematic at the time. In each case, the respective practices were or are considered normal inside the respective socio-cultural context, and thus, no men could be observed (Fleck & Kuzmics, 1985). Sociologists also found that whether or to what extent corrupt behavior comes to be critically reflected depends to a large extent on variables of education and social development. For example, the literature on the Italian Mafia reports that the well-known mafiosi were generally more or less illiterate (Fleck & Kuzmics, 1985). So even more than historical perspectives, sociological ones look at the social self-descriptions defining what is considered as being corrupt/unethical and what is not in different contexts, thus clearly treating corruption as a "phenomenon of perception" depending on the perspectives of those who analyze it (von Alemann 2005, 23).

Widely supporting these findings anthropological studies, in turn, are interested in micro level behavioral practices such as reciprocity with regard to their social function as central principles of human communication and thereby, as means of establishing relations of mutual trust (Mauss, 2002). A particular contribution of anthropology to the study of corruption can be seen in its focus on the perspectives of the acting individuals themselves. Moreover, studying the self-perception of "corrupt" actors shows to what extent social relations are a function of individual and collective sense-making, and how standards of measurement change with changing social ideals and identities. At any rate, anthropology teaches us that too strong or premature value judgments may prohibit an appropriate factual analysis of behavioral logics and the resulting social structures.

Last but not least psychological research has made important contributions to our understanding of corrupt behavior focusing, for example, on motives of bribe taking (Richter 1989). In view of a more complex and more integrative understanding of corruption, scholarship about the development of social perspective taking and moral judgment (Kohlberg, 1969; Rest, 1986; Chilton, 1988; Fein, 2012; and see below) is particularly interesting for two reasons. First, morality is at the basis of all our definitions of ethical and unethical/corrupt behavior, and second, this body of research is based on the idea of structurally different levels of complexity of possible ways to think about and act out morality. Even though this research has received considerable attention in the field of

behavioral ethics (Treviño, Weaver & Reynolds, 2006), we claim that it is sometimes insufficiently understood and so far not been systematically exploited and utilized. This is especially the case for Kohlberg's and others' finding according to which the majority of the adult population in western countries is functioning at the conventional levels of moral judgment while more principled, post-conventional structures of reasoning are empirically rare (Kohlberg, 1991, for an overview of Kohlberg's stages, see table 1 below). What does this mean for behavioral ethics in general and for corruption in organizations in particular?

» CORRUPTION VIEWED BY BEHAVIORAL ETHICS AND ORGANIZATION STUDIES: SOME OF THE MOST FREQUENT ANALYTICAL FRAMES, PATTERNS AND LINES OF ARGUMENTATION

Interestingly, academic economics have hardly perceived corruption as a problem for a long time. Due to the important role of neoliberal combined with rational choice perspectives in mainstream economics, deregulation tends to be generally appreciated by large parts of economic discourse while the merits of public regulation of markets have received much less attention. So while governmental and bureaucratic interventions into the "free play" of market mechanisms are often criticized, some economists have even hailed material incentives undermining state bureaucracies, for example in the case of the Soviet *Nomenklatura* (von Alemann 2005, 23). However, political economists like Susan Rose-Ackerman (2005) have meanwhile observed a growing readiness of business itself to accept broader ethical responsibilities. Also has ethics become an important issue in academic economic literature (White, 2009; Ulrich, 2008; Young, 1997; Sen, 1987), as well as in business ethics (Fisher/Lovell 2009, Waples et al. 2009) and (behavior focused) management studies (Kuhn/Weibler 2012; Treviño et al. 2006). Rose-Ackerman herself has given clear accounts of the negative impacts of corruption from a common welfare perspective (1999 and 2005).

The following paragraphs briefly review a number of more or less randomly chosen articles on corruption in organizations which have been recently published in leading academic journals in the fields of organizational behavior (OB), behavioral ethics (BE) and management studies (MS) which can, to some extent, be considered as being inspired by behavioral economics. Since this choice is not exhaustive and therefore only partly representative with regard to the discourse on corruption in the fields mentioned above, let alone for academic economics in a more general sense, we do not claim to make statements about the overall state of discussion in the respective disciplines in either substantial, theoretical or methodological respects. Rather, our interest is to look at general structural patterns visible in the articles reviewed, and to point out typical modes and models of analysis and argumentation which we either find helpful or, inversely, problematic in view of a broader, more integrative understanding of corruption. More precisely, we will ask to what extent interdisciplinary horizons, namely the basic contributions and findings of the "classic" disciplines mentioned before are taken into account by the OB/BE/MS literature, and at what point shortcomings in the sense of disciplinary reductionisms can be observed which could be overcome by a more integrative perspective as proposed elsewhere (Fein & Weibler, this issue). For systematic reasons we will limit this discussion to four of the most frequently found aspects indicative of typical patterns of analysis and argumentation in the OB/BE/MS literature, trying to offer a critical review with regard to the questions mentioned above.

In the following section, we first report and document our general observation that OB/BE/MS authors mostly do subscribe to the overall idea that the complexity of corruption can best be dealt with by using broader, i.e. more complex perspectives on the issue which is mostly understood as the challenge to integrate as many relevant aspects as necessary, and/or possible. However, we also detect that those general calls for theoretical and methodological integration and contextualization are often insufficiently met by authors themselves. This can be shown with respect to at least three aspects which we consider problematic: First, we found that moral and ethical judgments tend to remain insufficiently reflected and contextualized in most of the literature reviewed here. Second, actors' perspectives are often insufficiently taken into account, while rational choice presuppositions often remain insufficiently questioned. And third, disciplinary reductionisms are frequently visible in the shape of rather simplistic strategies of argumentation based on linear concepts of causation which go counter to the calls for contextualization cited in the first sub-section below.

Table 1. Correspondence of Stage Models (Kohlberg – MHC)

MHC stages			Kohlberg stages of moral development	
15	Cross-paradigmatic	(7)	(hypothetical)	
14	Paradigmatic	6	Universal ethical principles	Post-conventional morality
13	Meta-systematic	5	Social contract (may conflict with moral principles)	
12	Systematic	4	Authority and social-order maintaining, law and order	
11	Formal	3/4		Conventional morality
10	Abstract	3	Social expectations, interpersonal accord and conformity, good boy/girl	
9	Concrete	2/3		
8	Primary	2	Exchange, self-interest, what's in it for me?	Pre-conventional morality
7	Pre-operational	1/2		
6	Sentential	1	Obedience and punishment	
5	Nominal	0/1		
4	Sensory-motor	0		
3	Circular sensory-motor	-1/0		n.a
2	Sensory or motor	-1		
1	Automatic	-1/-2		
0	Calculatory	-		

Note. This table has been adapted from Commons & Sonnert 1994 and Tuladhar and Commons, 2014

» BROADER PERSPECTIVES AND CONTEXTUAL SENSITIVITY AS EXPLICIT GOALS OF CORRUPTION RESEARCH

To begin with, calls for a more interdisciplinary or even a "holistic view" of corruption can be found in a large number of publications. In fact, at first glance, most of the central requirements reported above (1. acknowledging historical and cultural contingencies of both corruption itself and our way of evaluating it, 2. acknowledging the difference between actors' perspectives and cultural norms, 3. acknowledging the fact that different actors tend to act on the basis of different cognitive, moral, motivational and other predispositions) are principally taken into account by authors writing about corruption from BE/OB/MS perspectives. But how are these calls framed and how do authors conceptualize their respective ideas about a truly integrative outlook on corruption?

As a rule, researchers stress the complexity of the phenomenon and therefore also call for epistemological complexity and theoretical integration (Lange, 2008). The latter are mostly understood as research designs which include multiple aspects and dimensions of corruption. Sometimes authors also make explicit calls for a contextualized view of corrupt actions and corruption as a social phenomenon in general. To name just a few:

» Masoud Shadnam and Thomas B. Lawrence (2011), focusing on ethical discourse and decision making in organizations, stress that "morality in organizations is embedded in nested systems of individuals, organizations and moral communities". Since they conceive of morality as "neither personal nor universal, but [as] always situated in a specific social and historical context", they claim that individual and organizational factors must not be regarded as standing in isolation from one another but that they rather have to be treated as interdependent. Moreover, they see ethical discourse and decision making in organizations as being "significantly influenced by a broad set of mechanisms and flows that connect moral communities, organizations and individuals". In particular, Shadnam & Lawrence urge for "thick descriptions" based on constructivism, i.e. for a more systematic inclusion of social and cultural contexts, as well as of methods and perspectives able to provide access to individual understandings of organization members' own behaviors.

» Similarly, Tanja Rabl (2011) in her piece on situational influences on corruption in organizations stresses the interdependence of different factors influencing corrupt behavior which she conceives of as the result of a "complex interplay of motivations, volitions, emotions, and cognitions in an individual's decision making process". Deploring that "there is little research focusing on the corrupt actors themselves", Rabl emphasizes the "relevance of all the person-related psychological components determining an individual's behavior".

» Focusing on the influence of administrative structures on corrupt behavior, Patrick von Maravic (2007a) criticizes that "conventional analysis of corruption ignores cultural dynamics and norms". To remedy this shortcoming, von Maravic himself suggests combining institutional and behavioral perspectives. He therefore bases his analysis of decentralized corruption in German municipalities on the theory of Actor-Centered Institutionalism (Scharpf, 1997).

» An explicit call for a "holistic perspective" on corruption is voiced by Yadong Luo (2004) who claims to deliver such a view by combining micro and macro-level perspectives, and considers, amongst others, aspects such as organizational design, task and institutional environments, organizational behaviors and anti-corruption practices (for a critique of Shadnam & Lawrence, Rabl, von Maravic and Luo see Fein & Weibler, this issue).

» Finally, among the contributions reviewed here, one of the most far-reaching and encompassing urges for a broader systems view has been pronounced by Ashforth, Gioia, Robinson & Treviño (2008). In their "Introduction to the Special Topic Forum: Re-Viewing organizational corruption", they propose to view "corruption in organizational life as a systemic and synergistic phenomenon". Focusing on the interrelations between "multiple perspectives and bodies of literature that can be brought to bear on the phenomenon", they hold that interdisciplinary research should consider psychological, sociological, cultural, economic, and political factors. According to Ashforth et al., the complexity of corruption can only be grasped if systemic perspectives consider both formal and informal, in other words if they also consider processes, behavior, ethical standards, and cognitive moral development. Moreover, these autors claim that there is "much need for conceptual work that is integrative, interactionist, and processual in nature", connecting micro, macro, wide, long and deep view: "We need a considerably more holistic or dynamic understanding regarding the interplay of environmental, organizational, and individual forces, i.e. a more macro view – to help us understand the etiology and evolution of corruption". What's more, Ashforth et al. assert a "necessity for a substantial shift in our way of thinking about both organization and society" the dimension of which they frame as a need for a "next wave of societal-level thinking" which also "considers the wider societal and even global implications of our actions".

To sum up, the authors cited not only perceive corruption as a challenging, complex phenomenon in the sense that they acknowledge multiple interrelations of structural/institutional/organizational with personal/motivational/behavioral aspects of corruption, as well as with its social, cultural and (to a lesser extent) historical dimensions. This complexity is also considered as a major challenge to corruption research. Anyhow, despite these general acknowledgements, shortcomings in view of meeting the challenge of this complexity analytically can be observed in many of the reviewed publications. This might be due to disciplinary biases in either of the fields considered here (organizational behavior, behavioral ethics and management). Below, we will report and discuss three of those shortcomings in more detail as a basis for proposing a more integrative explanatory framework of corruption in a second step (see Fein & Weibler, this issue).

Insufficiently reflected and/or contextualized value statements

The first problematic, yet frequently found structural characteristic of many OB/BE/MS articles on corruption, is that even though the influence of social, situational and cultural contexts on both corrupt behavior and on its evaluation is generally acknowledged, scholarship from those fields often does not reflect, question and/or contextualize its own moral and/or ethical standards and evaluation criteria in any systematic way. How does this relate to the general observation reported above?

We have just cited a number of statements arguing that cultural dimensions of corruption such as differing ethical standards in different social contexts and moral communities, as well as cognitive differences among individuals had to be included into our understanding of the phenomenon. This implies, first, that to behave (un)ethically or corrupt means different things in different contexts and for different people/social actors, and, as cultural historians have shown (see above), that it has also meant different things in different times. (This does not mean that what is considered as moral is or was completely contingent, see Fein & Weibler, this issue.) This insight implies, second, that standards and categories of evaluation need to be reflected, explained, contextualized and possibly also to be justified. Third, it implies that our own (personal and scientific) notions of moral and ethical behavior are themselves equally subject to and indicative of particular, often implicit sets of norms and values. In the following paragraph, we give a few examples of insufficiently reflected and contextualized value judgments with regard to ethics and morality in OB/BE/MS literature on corruption. However, the fact that we criticize this shortcoming in a number of cases does not mean that the respective publications do not otherwise make valuable contributions to our understanding of corruption. Moreover, the examples from literature on corruption cited in this article are merely intended to illustrate our respective claims and observations. We do not intend to thereby give comprehensive evaluations of the publications cited in any more general sense.

In this regard, as a first example, the CfP for a Special Issue on "Unethical Behavior" by a journal like *Organizational Behavior and Human Decision Processes* can be interpreted somehow ambivalently. While calling for a critical, refreshing reorientation of the field, the call also includes the tacit methodological presupposition of an "individual behavior that is subject to or judged according to *generally accepted* moral norms or social prescriptions". Yet, what is "generally accepted" differs largely between contexts, amongst others between different organizations, as well as between different actors within the same organization, as has been shown, amongst others, by studies from the fields of OB/BE/MS (Treviño, Weaver & Reynolds, 2006, Barmeyer & Davoine, 2011). Moreover, "general acceptability" appears to be a rather vage, if not questionable category, at least within a scientific context where the acceptability of corrupt behavior is one of the variables to be analyzed. We will therefore argue that the presupposition of a universal acceptance of certain norms is misleading when dealing with corruption, empirically and theoretically.

Nevertheless, our random literature review has discovered a surprisingly high number of similarly ambivalent statements made by corruption researchers from OB/BE/MS. For example:

» Luo (2004, see above), despite her supposedly "holistic" perspective, does not reflect and contextualize terms like "moral", "ethical" etc. used in her article. Summing up her description of the behavior of different "business types", she solely claims: "All these behaviors and underlying methods in response to task and institutional environments are illicit, immoral, unethical and illegal" (Luo, 2004).

» Studying corruption in financial institutions, Bertrand Venard and Mohamed Hanafi (2007) claim that "corruption is a cultural notion" (2007). However, they do not define the criteria for cultural variance with regard to corruption and how this relates to their study.

» In his article on corruption in organizational practice, Seraphim Voliotis (2011) does acknowledge that "organizations are embedded within a societal context, (that) widespread corrupt practices within the society are likely to be diffused to the organizations" (2011), and that thus, problematic practices like the "abuse (of authority) depend on the prevailing norms". He also stresses that therefore "each [type of corruption] needs to be treated distinctly". However, Voliotis does not problematize his notion of "deviant" as opposed to ethical behavior accordingly, but merely demands that organizations should offer ethical trainings to have their members comply with the organization's normative standards.

In fact, similar calls for compliance or conformism with organizational norms and standards are a pattern of argumentation we rather frequently found as an important aspect of the definition of "ethical behavior". The (rather circular) argument that if everyone of the organization's members behaved in the way implicitly or explicitly declared as ethical by the organization – and/or by the respective authors – there would be no problem, can thus be explained as a result of insufficiently reflected and/or contextualized notions of ethics and morality. This can be demonstrated with reference to the well-known stage model developed by Kohlberg (1981). For those calls for moral/ethical behavior suggest that making ethical decisions means conforming to particular rules. However, conforming to rules, according to Kohlberg's model (see table 1), is a conventional moral behavior (stage #4) as long as those rules are not questioned in the light of higher, more precisely universal principles (#5). Actually, many of the publications reviewed here make rather global demands on moral behavior, and in many of them, it is unclear whether by this they mean a conventional (Kohlberg stage #3 or #4) morality, i.e. that actors should behave in a certain "generally accepted" way, or a principled morality – which may at some point also imply questioning particular moral demands or conventions made by the organization if they are not compatible with higher principles. For example,

» Constant Beugré (2010), dealing with "deontic justice" as a way of preventing socialization into organizational corruption, suggests "to train employees in considering fairness as a moral obligation" without defining what "fairness", "morality" and "justice" mean and how their meanings possibly differ in different contexts and/or on different levels of moral development. Instead, the author gives two

(possibly incompatible) definitions of fairness both of which can be interpreted very differently depending on the level of personal/moral development of the respective actor. First, that "people seek fairness because it is the right thing to do". This can either be understood in a #5 principled way, i.e. as the demand to act according to higher principles that are independent of concrete social norms, or it can be a conventional moral reasoning in the sense that you just do things because you are expected to by some social group (#3) or norm (#4), and that you do not question what is considered to be the right thing to do by this group or norm. Second, Beugré argues that "a behavior is fair as long as it conforms to norms of moral obligation not only for oneself but also for others". This definition can equally be interpreted both in a conventional (conforming to norms) or in a principled way (moral obligation). So even though Beugré acknowledges that "corruption is both an act outside, behavioral perspective, E.F., J.W. and a state of being actors' consciousness perspective, E.F., J.W.", i.e. that both corruption and justice are relative in several respects, depending on inner/personal and outer/social or cultural factors, he does not differentiate and contextualize the moral ideals and concepts he uses accordingly. Rather, his poorly defined label "deontic" seems to qualify all actions as "ethical/moral" which correspond to his own, implicit principles. In this regard, note also his seemingly surprised remark that "otherwise decent people can end up engaging in questionable practices as a result of their immersion in, and socialization into, the social and cultural environment of a corrupt organization. We will come back to another problematic aspect of Beugré's argumentation after the next example.

» Kathie L. Pelletier and Michelle C. Bligh (2006) study the effectiveness of ethics programs in public sector organizations and the conditions "for an ethics program to be successful in educating employees about how to make decisions that are ethical" in the sense that they "serve the best interest of their customers and stakeholders". But they do not discuss what "morally/ethically appropriate" means if those interests conflict with other interests, as it is usually the case in real life moral dilemmas. While proposing that "ethical decision making is the process whereby individuals use their moral base to determine whether a certain situation or issue is right or wrong", the nature of this "base" is not specified, nor how its use may differ between individuals in various contexts and what this implies for the construction of ethics programs (which could, for example, also try to meet the needs of the different "clients" they wish to "serve" in a more flexible way). Even though the authors stress the "importance of informal norms on ethical behavior", the focus of their analysis is primarily on outcome effectiveness, i.e. on the effectiveness of given ethic programs. This effectiveness is studied through its perception by the organization's members and, supposedly, by the extent to which the latter conform to the respective standards of ethical decision making.

» A conformist definition of ethics/morality is also proposed Hal Hershfield, Taya Cohen and Leigh Thompson. In their study on the influence of self conceptions on unethical behavior, they claim that ethical behavior is "what is acceptable to the larger community" (2012).

» An even stronger conformist stance is contained in Shadnam & Lawrence's paper (2011) when they propose "continuous surveillance or members' perception of continuous surveillance (along with enclosure)" as a "requirement for moral regulation to be effective". Otherwise, "if employment conditions undermine enclosure and/or work arrangements diminish the effectiveness of surveillance, moral collapse is more likely to occur", so they claim. In other words, if members' thinking and acting is not controlled and "regulated", Shadnam & Lawrence see little chance that they observe the organization's rules.

These examples not only show that recent publications on corruption often take certain moral and ethical standards for granted without systematically defining, explaining or contextualizing them. What we consider particularly problematic is that despite frequent references to Kohlberg's theory of the development of moral judgment capacities, it appears that Kohlberg's findings are not being adequately perceived and have not been systematically integrated by all authors who make similar claims or demands to ethical behavior. While Kohlberg's model clearly defines, frames and differentiates between five structurally different understandings of morality – and the logics of (moral) behavior connected to them – the use of vague concepts such as "general acceptability", "moral obligation", "moral responsibility", "fairness" and "conformism to norms" leaves unclear what level of morality and judgment is desired. This therefore sometimes leads to problematic conclusions.

While Kohlberg's theory stresses that in general, not only post-conventional, but all behavior is considered "moral" by the respective actors, i.e. in their respective understandings of morality, authors writing about corruption/unethical behavior from OB/BE/MS perspectives often seem to just refer to common language uses of those terms. In common language, "moral" and "ethical" generally appeal to a high level of reflexivity (and often also to intensive inner and/or public debates) in order to arrive at truly "moral/ethical solutions". In other words, common language uses both terms in an almost Habermasian (1991) sense which has often been identified with Kohlberg's highest level of morality, stage #6 (for a discussion see Commons & Sonnert, 1994), i.e. it simply means the highest ethical standards available.

Given that more precise definitions are often missing in the literature, authors implicitly seem to have in mind those rather high moral standards, which would correspond at least to Kohlberg's concept of principled morality (#5, since evidence for a stage #6 has been rare), when they urge for "morally" or "ethically appropriate" behavior (see for example Pelletier & Bligh, 2006). In fact, some authors explicitly subscribe to this ideal, such as Constant Beugré (2010) by stressing that higher levels of moral development result in more ethical decisions, and that therefore people at post-conventional levels of development should act as "deontic agents" until fairness has been "internalized" by other members of the organization. Similarly, Pelletier & Bligh (2006) call for ethical leadership and for the recruitment of "ethical individuals in key

leadership positions to foster an ethical culture" and "effective ethics program(s)". However, neither of them substantially differentiates between stage #4 and #5 moralities. Instead, Kohlberg's stages are rather treated as mere analytical "concepts" (Beugré 2010).

A more systematic theoretical integration of the dimension of moral development – which in our view scholarship on corruption cannot afford to ignore – would have to acknowledge three things: First, empirical findings by Kohlberg and others according to which stage #5 moral judgment capacities are very rare among the average adult population in western societies, and that we are therefore unlikely to find a lot of them in social and business organizations. According to Treviño, Weaver & Reynolds (2006), "fewer than 20% of American adults reach the principled level, stages 5 or 6 (…), where actions should be more consistent with moral thought (See Rest et al. 1999). Other sources see the average dispersion of stage 5 morality among adults in western countries is below 10 percent (Cook-Greuter, 2000; Kohlberg, 1981). Michael Commons (personal communication) even estimates that only 1.5 % of the population is meta-systematic (which in his Model of Hierarchical Complexity approximately corresponds to Kohlberg's stage #5, see table 1). Undifferentiated calls for principled morality, post-conventional deontic agents and ethical leadership therefore (to some extent) seem to mix up ideals with reality. A more realistic approach would rather expect to find a majority of conventional reasoners (#3 and #4) among organizations' members and consequently ask how those can best be motivated to act morally to the best of their current capacity. Bengres assertion that there was "no systematic research on the impact of moral development on allied concepts such as ethics" and that "corruption remains largely unexplored in the OB literature because of the lack of theory to guide empirical research" (2010) appears at least surprising.

Second, the nature and complexity of principled morality also has implications for efforts to actively "educate" people in this direction. In principled morality, there is no simple "right or wrong", since there is no more focus on external authorities that could be consulted (in contrast to relevant others in #3 or social norms/laws in #4). What counts on stage #5, according to Kohlberg, is the process of comparing and evaluating norms and to decide about their hierarchical status in relation to universal values. This is why higher, post-conventional levels of moral reasoning can hardly be acquired by simple education programs in, say, a weekend training course. Rather, developing individual judgment capacities usually takes years of practical exposure to and experience in complex ways of dealing with conflicts between competing moral norms (Kohlberg, 1981).

Finally, since moral reasoning in a principled way also includes solving conflicts between universal principles and other organizational norms and goals, high level ethics and morality can very easily run counter the latter and result in "deviant behavior" (from the organization's perspective) which,—for ethical reasons—does not conform to organizational goals and principles. This is probably why this aspect is hardly ever discussed in any of the articles reviewed.

On this basis, we assume that the demands for moral/ethical behavior cited above are actually just stage #4 claims for norm conformity – which, after all, would not only be enough for an organization to function properly, but would enforcedly be in the best interest of most organizations, unless organizations are really prepared to receive transformative feedback and critique from their members, and given the statistical average dispersion of conventional reasoning structures among the adult population.

To sum up, our impression with regard to the overall outlook and way in which concepts of ethics and morality are treated in many of the reviewed articles, is that ethical behavior tends to be regarded as a pre-conceived analytic variable rather than as an object of more detailed, qualitative study of itself. Besides the problem of insufficiently contextualized values and concepts, the misunderstandings of Kohlberg reported here might also be linked to a second frequently found shortcoming which, we suggest, is equally due to disciplinary bias.

Insufficiently reflected anthropological presuppositions

The second structural characteristic we found problematic in many OB/BE/MS publications on corruption in view of a more integrative, "holistic" understanding of the phenomenon, is that authors often argue on the basis of reductionist rationalistic anthropologies without either supporting their empirical validity nor discussing their theoretical and epistemological implications. In this respect, it is probably less interesting to give examples of rational choice inspired theoretical and empirical scholarship, "generally assuming an individualistic and rationally self-interested focus on fairness for the self", (see Chang & Lai 2002) and which, according to Treviño, Weaver & Reynolds (2006) has been the standard perspective until recently. Instead, we prefer to point at some typical self-contradictions within the literature concerning this matter and to discuss the bias and reductionism contained within similarly strong anthropological presuppositions. In a second step (Fein & Weibler, 2014), we will then re-interpret those views on the basis of a wider epistemological perspective, using the Model of Hierarchical Complexity (MHC) and other developmental models of structural complexity.

Though authors often stress the importance of balanced views and despite frequent calls for combining different perspectives, our general observation is that rationalist behaviorist approaches, to different degrees, still dominate most of the publications on corruption from the fields of OB/BE/MS reviewed here. More precisely, this means that (1) presumptions on human nature and on the nature of human behavior tend to be limited to the assumption of rather simplistic rationalistic action logics ("what's in it for me?"), (2) that at the same time, internal realms of actors' cognitions, motivations and volitions are rarely studied in more detail, let alone in an open, explorative fashion, while instead 3., behavior is mostly analyzed as influenced by external factors. In this respect, again, we found oxymorons in a number of publications. For example

» In his piece on corruption control, Donald Lange (2008) presents an interesting model of four different strategies of corruption control, each of which proposes a reaction to corruption from a different level of moral development, even though Lange does not discuss this connection. In fact, his model makes interesting suggestions for dealing with

the complexity of corruption. For why, we will come back to this later. Lange himself, however, conceives of the four control strategies he identified as mere "types of reaction" without looking at the structural relation between those types of reaction and the level of personal (cognitive/moral) development of the actors setting them up on the one hand and of those who are to be controlled by them on the other. Moreover, despite his warning to "be aware (of) implicit assumptions about human nature" and his efforts to look at the latter in a more differentiated way, Lange uses implicit generalizing anthropological assumptions himself, when he talks about the "prototypical employee" who's behaviors had to be "predicted". By analyzing corruption control systems as "external influences", Lange essentially treats behavior as a dependent variable, more precisely, as a variable depending primarily on the external influences he studied (which, of course, is not to say that those influences are completely irrelevant).

» Dealing with situational factors influencing corruption in organizations, Tanja Rabl, as reported above, does stress the interdependence of different such factors, but she nevertheless assumes rational behavior as the standard logic of reaction of individuals to whatever situational influence she considers to be relevant. So here again, behavior is perceived as a predictable variable, depending on quantifiable factors such as the size of bribes, time pressure and the degree of abstractness of the business code of the respective organization. Focusing exclusively on the relationship between these external variables and individual behavior, factors internal to the individual such as, for example, the level of complexity of their personal meaning making and/or moral development, which cause different individuals to react differently to the same external influences, are not taken into account. Even though, as reported above, Rabl (2001) deplores the lack of research on "person-based determinants of corruption", she does not make own efforts in this direction herself, but seems to assume that all individuals act according to the same "motivations, volitions, emotions and cognitions".

» The same critique applies to the structural logic of Patrick von Maravic's study on decentralized corruption mentioned above. By chosing the theory of actor centered institutionalism as a way of connecting two important perspectives, von Maravic (2007a) tries to overcome unidimensional perceptions of corruption. However, when combining institutionalist with rational choice perspectives, he equally limits his analysis to external influences on behavior. Even though he rightly criticizes "pure rational choice analysis" for neglecting certain "factors that have been considered important in the literature", his own strategy does not eliminate this blind spot. For when he studies the "influence of institutions on the individual set of preferences", von Maravic assumes that rational "risk calculation", in this case depending on the institutional environment", is the only relevant logic of individual behavior. Even if actors now "make their decisions within institutions" exerting various influences on them, it is apparently still theorized that the reactions of all individuals to the same influences are the same. They namely

consist of calculating the risk of being caught against the chances of getting advantages through corrupt behavior – a behavior typical of Kohlberg's stage #2 morality.

In contrast to those generalizing presumptions, developmental models studying the growing complexity of cognition, moral judgment, meaning making etc. such as those of Commons, Kohlberg, Kegan, Cook-Greuter and others, argue that rational risk calculation is the typical behavioral strategy of *one* particular developmental action logic, namely the formal stage (#11) in terms of the Model of Hierarchical Complexity (MHC, see table 1, for a detailed discussion see Ross & Commons, 2008 and Fein & Weibler, 2014), or the institutional or conscientious self in the models of Kegan and Cook-Greuter, but that it is not at all representative of other stages and their action logics (Commons, 2008; Ross & Commons 2008; Kegan, 1994; Cook-Greuter, 2000, see also Fein & Weibler, 2014). The term "action logic" has originally been introduced by Bill Torbert for describing stage dependent differences in the behavior of leaders. The definitions of stages are based on stage descriptions of the development of the self by Jane Loevinger and Susanne Cook-Greuter (see for example Torbert, 2000). Admittedly, a considerable number of adults in contemporary societies function on the basis of this action logic. However, given the well-established empirical basis of developmental theory (Modgil & Modgil, 1985), theoretical and explanatory models aiming at generalizable statements about human behavior, in our view, have to take into account that behavioral motivations differ significantly in relation to personal development, and therefore cannot be reduced to one uniform action logic mistaken as being able to explain the behavior of all humans.

Yet, to a lesser degree, the use of implicit rationalistic presumptions about human behavior is visible even in OB/BE/MS publications that do place a strong focus on inner, psychological aspects of the individual. Hal Hershfield, Taya Cohen and Leigh Thompson (2012), for example, study the influence of self conceptions on unethical behavior, but they still seem to have trouble going beyond traditional mental habits of their discipline, i.e. beyond the usual rational choice anthropology. Fcocusing on the degree to which corrupt tendencies are linked to the sense of continuity of the individual self over time, they argue that "one underlying cause of unethical conduct is a fundamental inability to project one's self into the future" and that "feeling disconnected from one's future self is intimately linked to unethical decision making". In other words, individuals with rather short horizons are more likely to give in to tempting offers and/or situations whereas those with a stronger sense of self continuity tend to resist similar temptations, i.e. behave (more) ethically. This hypothesis is completely in line with theories of moral and self-development, and Hershfield et al.'s findings are therefore not surprising. Since short horizons (because of their relatively low degree of complexity of perspective taking) are a typical feature of less advanced levels of cognitive and moral development, their relationship with corruption can easily be predicted from Kohlberg's and related theories even without empirical testing. However, in this structuralist perspective, short horizons are just a secondary effect of a more basic, primary fact, namely the structure of cognition (in its various dimensions).

Even though the authors explicitly refer to and take notice of "individual differences in cognitive moral development" and their effects in practical "ethical decision making and moral behavior", they do not link their findings systematically to the meta-frames made available by developmental theories, but rather treat the two categories (short horizons and continuity of self) as two among "many (other possible) determinants of unethical behavior", as if they were probing into completely unknown territory. So obviously, here again, the quality of developmental complexity theories as meta-models offering non-arbitrary hierarchical ordering principles not only for understanding human behavior, but also for evaluating theories making anthropological claims, is not fully appreciated.

This assumption is supported by the fact that Hershfield et al. treat developmental theories of the self (despite their empirical basis demonstrating that the structure of the self becomes more complex as it develops) just as one of several speculative *models* of the self, besides other models some of which claim, for example, that "people do not possess a continuous self over time" at all (Strotz, 1956, cited after Hershfield et al., 2012), and which thus clearly contradict empirical findings. On these grounds, in order to solve their theoretical dilemma, the authors suggest that an individual may have "multiple selves" which they see as types coming to determine individual actions in different situations one after another in a more or less random way. In other words: Theoretical integration comes down to the relativist assertion that each theory of the self is equally true, and that to choose between them is either arbitrary or impossible.

Nevertheless, the authors still seem to be looking for general laws and mechanisms "under what circumstances these alternate decision making selves emerge". Moreover, their research design, trying to achieve, or at least to study ethical behavior based on monetary gift certificates of $ 50, and thus, on the assumption of a rational choice motivational structure, shows that the rationalist bias is still influential, even though they admit that "it is fairly difficult to change future self-continuity with short-term interventions".

Finally, a similar perspective is visible in Zyglidopoulos, Fleming & Rothby's (2009) study on rationalization which, put in mhc terms, describes the action logic of the rational individual trying to "beat the system" (mhc #11; see Fein & Weibler, this issue) very well, but does not reflect the limits of this action logic in a broader sense. Again, the authors make generalizing claims with regard to human behavior such as that all humans tend to use rationalization strategies in order to defend and/or legitimize corrupt actions, without differentiating between behaviors motivated by different degrees of complexity of their action logics. While the statement above concerning rationalization is true for the formal stage which is able to coordinate two variables (own interest and the functioning of the system), and can thus be expected to have a minimal sense of wrongdoing, less complex action logics are unable to perceive the difference between their own interest and the requirements of a normative system. Hence, they would probably be less inclined to use rationalization. Moreover, Treviño, Weaver & Reynolds (2006) and others, in line with anthropological and historical research, have observed that "corrupt individuals (on those less developed levels) tend not to view themselves as corrupt". The perception of corruption is thus, to a large extent, a matter of

perspective. In contrast, more complex action logics might either use more sophisticated rationalization/justification strategies, but might as well – more probably – just refrain from corrupt actions altogether, because they see their own interest less separate from that of the overall system as a whole.

Again, even though the cognitive dimension is referred to by Zyglidopoulos et al., it seems to be treated merely as an open container for different kinds of thoughts and motivations rather than as a recordable precondition structuring all thinking and acting in a non-arbitrary, non-contingent way. So when the authors state that "what counts as corrupt is culturally and historically contingent", this claim (which is not systematically elaborated), is apparently merely used to imply that there is no way of explaining historical and cultural differences. In contrast to this (empirically insufficiently based) relativist position, we argue that cultural in combination with developmental studies have provided sufficient evidence to discard this relativist position in favor of a more complex, more multi-dimensional meta-perspective interested in the reasons for and the deeper causes of those cultural differences.

A rare exception among the articles on corruption considered here which comes closest to our own integrative vision (see Fein & Weibler, this issue) is the conclusion of the meta-study by Treviño, Weaver & Reynolds (2006). Since their contribution is itself a review of recent literature on behavioral ethics in organizations, it is perhaps more susceptible to the different aspects and dimensions studied by authors in the field, as well as to the complex interrelations between those dimensions, namely between thinking and acting, between individual and contextual factors, as well as to the fact that individuals behave differently in different contexts "based on their cognitive predispositions".

Treviño et al. not only clearly state the relation between self, moral identity and the respective action logics ("behavior, affect, and cognition are closely linked to self-identity" since "identity itself is formed through social cognition processes", cited after Bandura 1986). They also account for their changes in character and their changing influence on behavior in relation to developmentally acquired increases in complexity. As to the latter, Treviño (1986) has made clear that "the influences of contextual variables on decision making and behavior depend upon the individual's cognitive moral development, with those at the highest stages being less susceptible to contextual influences". In other words: Contextual factors lose their impact on individual action with increasing cognitive and self development, a fact that is hardly ever taken into account by the rest of the literature dealing with external influences on behavior which has been reviewed here.

Finally, Treviño et al. (2006) also stress the practical relevance of their statements: All of this "has clear implications for behavioral ethics in organizations" and therefore has to be taken into account in view of creating and shaping appropriate organizational motivational systems, they claim. Moreover, they consider a vital management task not only to invite organization's members to behave ethically, but to do this in a way that is adequate to the average structure of reasoning of the respective actors. Anyhow, they make clear that "leaders' moral reasoning is linked to their leadership style" which is why "organizations should consider using measures of moral reasoning to identify individuals for leadership

development or to assign them to leadership roles, particularly if they are going to lead group decision making about ambiguous ethical issues" – a proposition to which we subscribe and which is taken on in Fein & Weibler, this issue.

So how can the overall epistemological problem of the structural complexity of research designs, as well as of analytical perspectives in the literature reviewed here be summarized?

» CONCLUSION

This paper has argued that OB/BE/MS literature on corruption often fails to offer an adequate, i.e. multi-faceted and integrative understanding of the phenomenon, due to disciplinary constraints and traditions apparently inducing many researchers to take less-than-sufficiently complex views onto the phenomenon. This has been shown, first, with regard to often insufficiently reflected analytical terms and concepts in general and to a widespread lack of contextualization of value judgments in particular. Second, our claim has been illustrated by a widespread tendency to underrate out or ignore major aspects of the subjective dimension of behavior, namely by the reluctance to explore actors' empirical action

logics. Both of these aspects can be synthesized within a broader epistemological critique concerning the dominant strategies of analytical perspective taking. In this regard, we argue that a considerable number of contributions to current academic discourse on corruption in the fields mentioned above choose analytical perspectives that are narrower than necessary to understand the complexity of corruption adequately.

Moreover, these mainstream western notions of corrupt and/or unethical behavior (which are usually taken for granted as "generally accepted" in large parts of the Behavioral Ethics, Organizational Behavior and Management Studies literature, for example) are not at all representative for other than the present-time western context and are therefore not very helpful for dealing with corruption in many of those other contexts (see also Barmeyer & Davoine, 2011), at least according to our preliminary estimate. In this regard, it has to be asked to what extent and how these mainstream notions can and/or have to be contextualized – or at least to be made more explicit – in order to be able to make more generally valid claims about the phenomena in question (for answers see Fein & Weibler, this issue). ■

REFERENCES

Aguilera, R.V., & Vadera, A.K. (2008). The dark side of authority: Antecedents, mechanisms, and outcomes of organizational corruption. *Journal of Business Ethics*, 77(4), 431–449. doi: 10.1007/s10551-007-9358-8

Anand, V., Ashforth, B. E., & Joshi, M. (2004). Business as usual: The acceptance and perpetuation of corruption in organizations. *Academy of Management Executive*, 18(2), 39-53. doi: 10.5465/AME.2004.13837437

Arlacchi, P. (1989). *Mafiose Ethik und der Geist des Kapitalismus: Die unternehmerische Mafia*. Frankfurt/Main: Cooperative-Verlag.

Ashforth, B. E., & Anand, V. 2003. The normalization of corruption in organizations. In R. M. Kramer & B. M. Staw (Eds.), *Research in organizational behavior* (pp. 1-52), Oxford: Elsevier.

Ashforth, B. E., Goia, D. A., Robinson, S. L., Trevino, L. K. (2008). Introduction to special topic forum: Re-viewing organizational corruption. *Academy of Management Review*, 33(3), 670–684. doi: 10.5465/AMR.2008.3246

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, New Jersey: Prentice Hall.

Barmeyer, C. I., Davoine, E. (2011). Die implementierung wertfundierter nordamerikanischer verhaltenskodices in deutschen und französischen tochtergesellschaften. Eine vergleichende fallstudie. *Zeitschrift für Personalforschung, 25*(1), 5-27.

Beugré, C. D. (2010). Resistance to socialization into organizational corruption: A model of deontic justice. *Journal of Business Psychology*, 25(3), 533–541. doi: 10.1007/s10869-010-9176-3

Chang, J. J., & Lai, C. C. (2002). Is the efficiency wage efficient? The social norm and organizational corruption. *Scandinavian Journal of Economics 104*(1), 27-47. doi: 10.1111/1467-9442.00270

Chilton, S. (1988). *Defining political development*. Boulder, Colorado: L. Rienner Publishers.

Commons, M. L. (2008). Introduction to the model of hierarchical complexity and its relationship to postformal action. *World Futures 64*(5-7), 305–320. doi: 10.1080/02604020802301105

Commons, M.L., & Sonnert, G. (1994). Society and the highest stages of moral development. *Politics and the Individual*, 4(1), 31-55.

Cook-Greuter, S. (2000). Mature ego development: A gateway to ego transcendence? *Journal of Adult Development*, 7(4), 227-240. doi:10.1023/A:1009511411421

Engels, J. I. (2010). Politische korruption und modernisierungsprozesse: Thesen zur signifikanz der korruptionskommunikation in der westlichen moderne. [Political corruption and modernization] In Grüne, N. & Slanička, S. (Eds.), *Korruption: Historische annäherungen* (pp. 35-54). Göttingen: Vandenhoeck & Ruprecht.

Engels, J. I., Fahrmeir, A., & Nützenadel, A. (2009) (Eds.), *Geld, Geschenke, Politik: Korruption im neuzeitlichen Europa*. München: Oldenburg Verlag.

Fein, E. (2012). Kognition und politische Kultur. Sozialpsychologische perspektiven in der diskursforschung am beispiel von patronage, klientelismus und korruption in Russland. Cognition and political culture. Adult development perspectives on patronage, clientelism and corruption in Russia] In Oberthuer, J., Rosa, H. (Eds.), *Normen- und Wertbegriffe in Europa im Wandel* (in print).

Fein, E., & Weibler, J. (2014). Cognitive basis for corruption and attitudes towards corruption viewed from a structuralist adult developmental meta-perspective. Behavioral Development Bulletin, 19(3).

Fisher, C., & Lovell, A. (2009). *Business ethics and values: Individual, corporate and international perspectives* (3. ed.) Pearson Education Prentice Hall: Englewood Cliffs.

Fleck, C. & Kuzmics, H (Eds.) (1985). *Korruption: Zur Soziologie nicht immer abweichenden Verhaltens*. [Corruption: Sociology of not always deviant behavior] Königstein/Ts.: Athenäum.

Grüne, N., & Slanička, S. (Eds.), *Korruption: Historische annäherungen an eine grundfigur politischer kommunikation*. [Corruption: Historical approaches] Göttingen: Vandenhoeck & Ruprecht.

Habermas, J. (1991). *Moral consciousness and communicative action*. Cambridge: MIT Press.

Hershfield, H., Cohen, C., & Thompson, L. (2012). Short horizons and tempting situations: Lack of continuity to our future selves leads to unethical decision making and behavior. *Organizational Behavior and Human Processes, 117*(2), 298-310. doi: 10.1016/j.obhdp.2011.11.002

Hutchinson, M., Vickers, M.M., Wilkes, L., & Jackson, D. (2009). The worse you behave, the more you seem to be rewarded": Bullying in nursing as organizational corruption. *Employee Responsibilities and Rights Journal, 21*(3), 213—229. doi:10.1007/s10672-009-9100-z

Jain, A. K. (2001). Corruption: A review. *Journal of Economic Surveys, 15*(1), 71-121. doi: 10.1111/1467-6419.00133

Johnston, M. (2005). Keeping the answers, changing the questions: Corruption definitions revisited. In von Alemann, U. (Ed.), *Dimensionen politischer korruption. Beiträge zum stand der internationalen forschung* (pp. 61-76). Wiesbaden: VS.

Kegan, R. (1994). *In over our heads: The mental demands of modern life*. Cambridge, Massachusetts: Harvard University Press.

Kohlberg, L. (1981). *Essays on moral development, Vol. I: The philosophy of moral development*. San Francisco, California: Harper & Row.

Kohlberg, L. (1969). Stage and sequence. The cognitive developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory* (pp. 347-480). Chicago: Rand McNally.

Kuhn, T., & Weibler, J. (2012). *Führungsethik in organisationen*. Stuttgart: Kohlhammer.

Lange, D. (2009). A multidimensional conceptualization of organizational corruption control. *Academy of Management Review, 33*(3), 710—729. doi: 10.5465/AMR.2008.32465742

Luo, Y. (2004). An organizational perspective of corruption. *Management and Organization Review*, *1*(1), 119—154. doi: 10.1111/j.1740-8784.2004.00006.x

Maravic, P. V. (2007). Decentralized corruption in Germany. *Public Administration Quarterly*, *30*(4), 440-478.

Maravic, P. V. (2007b). Public management reform and corruption: Conceptualizing the unintended consequences. *Administration and Public Management Review, 8*, 126-148.

Mauss, M. (2002). *The gift: The form and reason for exchange in archaic societies*. London: Routledge.

Modgil, S., & Modgil, C. (Eds.) (1985). *Lawrence Kohlberg: Consensus and controversy*. Philadelphia and London: Falmer Press.

Pelletier, K.L. & Bligh, M.C. (2006). Rebounding from corruption: Perceptions of ethics program effectiveness in a public sector organization. *Journal of Business Ethics, 67*(4), 359—374. doi: 10.1007/s10551-006-9027-3

Plumpe, W. (2009). Korruption: Annäherungen an ein historisches und gesellschaftliches Phänomen. [Corruption, viewed as a historical and social phenomenon] In Engels, J.I., Fahrmeir, A., & Nützenadel, A. (Eds.), *Geld, Geschenke, Politik: Korruption im neuzeitlichen Europa* (pp. 19-48). Munich: Oldenburg Verlag.

Rabl, T. (2011). The impact of situational influences on corruption in organizations. *Journal of Business Ethics*, *100*(1), 85—101. doi: 10.1007/s10551-011-0768-2

Rest, J. R. (1986). *Moral development: advances in research and theory*. New York: Praeger.

Richter, H. E. (1989). *Die hohe kunst der korruption: Erkenntnisse eines politik-beraters*. Hamburg: Heyne Verlag.

Rose-Ackerman, S. (1999). *Corruption and government: Causes, consequences, and reform*. Cambridge: Cambridge University Press.

Rose-Ackerman, S. (2005). Groß angelegte korruption und ethik in der globalen wirtschaft. In von Alemann, Ulrich (Ed.), *Dimensionen politischer korruption. Beiträge zum stand der internationalen forschung* (pp. 195-229). Wiesbaden: VS.

Scharpf, F. (1997). *Games real actors play: Actor-centered institutionalism in policy research*. Boulder, Colorado: Westview Print.

Scott, C. (1972). *Comparative political corruption*. Englewood Cliffs: Prentice Hall.

Sen, A. (1987). *On ethics and economics*. Malden, Massachusetts: Blackwell.

Senturia, J. A. (1930). Corruption, political. In *Encyclopedia of the Social Sciences, Vol. 4* (pp. 448-452). New York.

Shadnam, M., & Lawrence, T. B. (2011). Understanding widespread misconduct in organizations: An institutional theory of moral collapse. *Business Ethics Quarterly, 21*(3), 379-407.

Torbert, W. R. (2000). A developmental approach to social science: A model for analyzing Charles Alexander's scientific contributions. *Journal of Adult Development, Vol. 7*(4), 255-267. doi: 10.1023/A:1009515512330

Torbert, W. R. et al. (2004). Action inquiry: *The secret of timely and transforming leadership*. San Francisco: Berrett-Koehler Publishers.

Treviño, L. K., Weaver, G. R., & Reynolds, S. R. (2006). Behavioral ethics in organizations: a review. *Journal of Management* 2006, *32*(6), 951-990. doi: 10.1177/0149206306294258

Ulrich, P. (2008). *Integrative economic ethics: Foundations of a civilized market economy.* Cambridge: Cambridge University Press.

Venard, B., & Hanafi, M. (2008). Organizational isomorphism and corruption in financial institutions: Empirical research in emerging countries. *Journal of Business Ethics, 81*(2), 481-498. doi: 10.1007/s10551-007-9519-9

Voliotis, S. (2011). Abuse of ministerial authority, systemic perjury, and obstruction of justice: Corruption in the shadows of organizational practice. *Journal of Business Ethics, 102*(4), 537-562. doi: 10.1007/s10551-011-0826-9

Volkov, V. (2000). Patrimonialism versus rational bueraucracy : On the historical relativity of corruption. In Lovell, S. et al (Eds.), *Bribery and Blat in Russia*: *Negotiating reciprocity from the early modern period to the 1990s* (pp. 35-47). London: Palgrave Macmillan.

von Alemann, U. (2005). Politische korruption: Ein wegweiser zum stand der forschung. In von Alemann, U. (Ed.), *Dimensionen politischer korruption*: *Beiträge zum stand der internationalen forschung* (pp. 13-49). Wiesbaden: VS.

Voslensky, M.S. (1987). *Nomenklatura*: *The* Soviet *ruling class*. New York: Doubleday.

Waples, E. P., Antes, A. L., Murphy, S. T., Connelly, S., & Mumford, M.D. (2009). A meta-analytic investigation of business ethics instruction. *Journal of Business Ethics. 87*(1), 133-151. doi: 10.1007/s10551-008-9875-0

White, M.D. (2009). Introduction to ethics and economics. *Review of Social Economy, 67*(1), 1-2. doi: 10.1080/00346760802460693

Young, J.T. (1997). *Economics as a moral science.* Cheltenham: Edward Elgar.

Zyglidopoulos, S. C., Fleming, P. J., & Rothenberg, S. (2009). Rationalization, overcompensation and the escalation of corruption in organizations. *Journal of Business Ethics, 84*(1), 65—73. doi: 10.1007/s10551-008-9685-4

Cognitive basis for corruption and attitudes towards corruption in organizations viewed from a structuralist adult developmental meta-perspective

Elke Fein and Jürgen Weibler

University of Hagen, Germany

ABSTRACT

The paper focuses on corruption and attitudes towards corruption in organizations. It proposes an interdisciplinary framework for reassessing them. It is argued that an integrative theoretical and analytical framework based on the Model of Hierarchical Complexity (MHC) can provide new insights on corruption. Furthermore the proposed framework offers new theoretical horizons for understanding and evaluating public and scientific discourses on corruption. This approach compensates for frequent shortcomings and disciplinary reductionisms in large parts of the social science literature on corruption. It can thus offer a substantially new outlook on the field of behavioral ethics in organizations based on a meta-systematic theory integration.

KEYWORDS: corruption, unethical behavior, behavioral ethics, adult development, complexity, action logics, stages, moral development, theoretical integration, contextualization

The present paper is motivated by a number of surprising, puzzling observations about present theorizing about corruption – and the desire to resolve them. In a nutshell, our primary theoretical motivation is to account for the plurality of strikingly different and sometimes even contradictory perspectives that important social sciences disciplines currently dealing with corruption take onto the phenomenon, and to propose a theoretical framework which is comprehensive enough to bridge those differences. Moreover, we claim that the Model of Hierarchical Complexity (MHC, see below) is able to integrate the contributions that each disciplinary perspective has to make, in order to provide a more complex and more differentiated understanding of corruption. The remainder of this introduction will explain our observations, our interpretation, our motivation and the approach that resulted from those observations and interpretations in some more detail.

Behavioral ethics in general, as well as corruption (as an important form of unethical behavior) in particular have gained increasing public and scholarly interest during the last decades. The number of publications on the topic has grown rapidly, not only in the field of behavioral ethics (Treviño, Weaver & Reynolds, 2006), but also in almost all major social sciences (Grüne & Slanička, 2010; Heidenheimer & Johnston, 2007; Jain, 2001). This

influx of attention, we suggest, is a product and an expression of two general cultural trends in most western societies. The two cultural trends are:

1. A general cultural and political sensitization for ethical questions and,

2. An increasing self-reflexivity of both sciences and societies in many respects.

However, depending on which disciplinary lens is being used, the abundant literature on corruption varies greatly in perspective, focus and evaluation of the phenomena. Luo (2004) even speaks of "fundamentally different paradigmatic perspectives" through which each discipline dealing with corruption looks at the issue. Besides the field of Behavioral Ethics, which is strongly influenced by psychological research, corruption in (and outside) organizations has so far been studied from historical, sociological and anthropological perspectives, to name only the most important ones (see Fein & Weibler, this issue). At the same time, convincing interdisciplinary syntheses remain rare. Overarching analytical categories are almost completely missing. This has resulted in disciplinary reductionisms and widespread theoretical and analytical shortcomings. Some of this can also be spotted within the field of behavioral ethics in organizations and its outlooks on corruption. The present article attempts to account for some of these shortcomings by questioning how these problems can be solved and how the differences and contradictions between and within the disciplinary approaches can eventually be bridged.

Fein & Weibler (2014) reviewed literature on corruption and found substantially different understandings of corruption not only in different social contexts and in different times, but also in different scientific cultures. They pointed out that a considerable number of contributions to current academic discourse on corruption choose analytical perspectives that are narrower than necessary to understand the complexity of corruption adequately. The following paragraph summarizes the most important results of our (non-exhaustive) literature review with respect to the structural complexity and the degree of self-reflexivity of the articles considered.

Paradigmatic shortcomings in organizational behavior (OB), behavioral ethics (BE) and management studies (MS) literature

Note that the principal criticisms mentioned below are not substantially new and have been raised many times, especially with respect to some parts of quantitative research, which is why we limit ourselves to rather short statements in this regard.

a) Constructing plain variables out of complex phenomena. By this we mean approaches which build single, and often too simple variables out of complex real life phenomena, as for example to describe and quantify "situational" (Rabl 2011) or "institutional" influences (von Maravic, 2007a) on corrupt behavior or that of the "overall working atmosphere" (Chang & Lai, 2002). Of course, factors such as the size of bribes, time pressure and the degree of abstractness of the business code can be isolated and quantified (as in Rabl, 2011), but by making similar choices, other, potentially, equally or even more relevant factors are left out of the analysis, often for practical reasons. This critique is related to the next one.

b) Too small number of variables. Generally, human behavior is influenced by complex webs of (internal and external) factors and by the dynamics of their interrelations. Therefore, at least from a meta-systematic perspective (see below), it appears highly insufficient to reduce the explanation of behavior to, say, two single variables. However, singling out and quantifying individual factors from a given context and making statements in the sense that one was able to explain the respective behavior on these grounds is still a widespread scientific practice. However, in our view, this academic habit is more instructive than the respective researcher's cognitive/analytic perspective than it actually illuminates the object under research.

c) Overexaggeration of linear assumptions of causality between variables. After variables have been identified, operationalized and their statistical values quantified, research designs often use quite simple, linear descriptions of (at least indirectly assumed) causal relations between the variables (Venard & Hanafi, 2007; Beugré, 2010; Pelletier & Bligh, 2006; Lange, 2009; Rabl, 2011; Aguilera & Vadera, 2008; Shadnam & Lawrence, 2011). Even though authors might acknowledge (on the basis of thorough statistical analysis) that causal relations between variables are complex and/or not as straightforwardly discernable as hypothesized before, and even though they might be conscious of the fact that the explanatory scope of their findings is limited due to the constraints mentioned above, linear causation still seems to have a great appeal to many researchers. This might be because linear formulations of causation between

variables always implicitly create the impression of expressing quasi-natural laws – and thus being more "objective" by reporting "pure scientific truths" with regard to the matter in question – while more complex, qualitative explorations into the respective context or phenomenon are often left open "to future research".

d) Neglecting subjective dimensions. As already mentioned above, (quantitative) research interested in supposedly clear causal relationships often tends to leave out those dimensions of the respective phenomenon which are "difficult to access" or at least difficult to operationalize and/or quantify. Very often, this argument applies to the empirical reality of individual actors' motivations, i.e. the latter are either left out of the analysis or reduced to simplified assumptions instead of being explored empirically (Luo, 2004; Shadnam & Lawrence, 2011). In our view, it is clear that on these grounds, no "holistic view" of human behavior is possible.

e) Missing integration of dimensions and perspectives. Finally, as reported in Fein & Weibler (this issue), a number of authors do stress the interdependence between various dimensions of corruption and, therefore, the importance of integrating perspectives. Some even make more concrete suggestions for future research in this regard. However, ultimately, only a small portion of the articles reviewed actually come close to meeting these demands themselves. Of course, a lack of interdisciplinarity can hardly be reproached to single research papers from single disciplines. Anyhow, since the fields considered here and, in more detail, in Fein & Weibler (this issue) are themselves already more or less interdisciplinary endeavors, we consider it legitimate to address certain reductionisms, as we see them. So when giving examples in the previous paragraph, we thereby do not claim that the authors working and writing within the respective traditions are doing "bad research", but we do claim that the underlying patterns of thought are often less complex than desirable. Moreover, our argument is that in combination, all of the shortcomings listed here can be seen as results of particular epistemological paradigms which correspond to particular levels of complexity of analytic perspective taking which, in our view, are not sufficient in view of an encompassing understanding and explanation of corruption, because they are too narrow to grasp its complexity and dynamics. In order to go beyond these limitations, we wish to introduce the Model of Hierarchical Complexity as a tool for analyzing and measuring the complexity of both empirical phenomena of corruption and our way of understanding and conceiving of those phenomena (attitudes towards corruption).

Intention of the current paper

To counter the problems and shortcomings mentioned above and, in more detail, by Fein & Weibler (2014), we offer a synergistic, interdisciplinary framework based on, on the one hand, questions and findings from various other social sciences, and, on the other hand, a theoretical explanatory model suggesting abstract criteria for analyzing and evaluating corrupt behavior itself and the discourse dealing with it. Focusing on the inherent patterns of structural complexity of both physical and discursive behavior, we thereby expand on and tie up with Ross & Commons' basic paper on political development (2008), as well as with Treviño,

Weaver, and Reynolds' (2006) claim that cognitive developmental dimensions have "clear implications for behavioral ethics in organizations". Moreover, we argue that adult developmental perspectives, namely the Model of Hierarchical Complexity (MHC; Commons 2009; Commons, Sinnott, Richards, Armon 1989; Commons, Armon, Kohlberg, Richards, Grotzer, & Sinnott, 1990; Levinson, 1986; Marchand 2002) can provide content-free analytic tools and thus introduce more complexity and analytical rigor into the study of corruption in organizations.

By adult development research, we mean different strands of structuralist scholarship on the development of cognitive and other mental capacities and competences inspired by pioneers like Baldwin, Mead, Piaget and others, claiming – and providing empirical evidence for the fact that the development of cognition and other aspects of mind does not stop after adolescence, but can continue in adulthood, even though in many cases it does not for structural and contextual reasons. This has strong, yet still largely underestimated consequences not only for social, political and economic life in general, but also for human behavior in organizations and, on this basis, the functioning of organizations, including states, on micro, meso, and macro levels.

In this paper, we focus on the *Model of Hierarchical Complexity (MHC)*, a mathematics-based general stage model developed since the late 1970's and early 1980's by Michael L. Commons, Francis Richards and others (Commons, 2008; Commons & Richards, 1984a; Commons & Richards, 1984b; Ross & Commons, 2008). By its nature as a content-free general model, the MHC accounts for the developmental content-based stage progressions in over 33 currently known developmental models, among them that of moral judgment by Lawrence Kohlberg which is widely acknowledged and discussed within the field of behavioral ethics (see for example Treviño, Weaver & Reynolds' review, 2006). Using the MHC, structurally different forms of corruption can be identified according to their inherent level of complexity. Furthermore, the structural complexity of ways of dealing with and of theorizing about corruption becomes visible so that theoretical and practical anti-corruption activities can be better adjusted to the nature of the problems.

We make two claims on this basis. First is that explaining corrupt behavior in terms of the development of structural complexity can lead to a better understanding of where ethical and unethical behavior as discussed in the literature on organizational behavior, behavioral ethics and management studies (OB/BE/MS). We can explain why behaviors themselves and attitudes towards corruption differ tremendously in different (scientific and organizational) cultures and contexts, and why the dominant, supposedly "generally accepted" notions of (un-) ethical or corrupt behavior in mainstream western culture and science must not be generalized. We can also explain why their unconscious projection on other contexts may lead to counterproductive results. Second, we claim that re-evaluating the plurality of current theorizing about corruption according to its inherent patterns of structural complexity can help to overcome disciplinary reductionisms and thus push behavioral ethics research into new meta-theoretical or paradigmatic insights. Doing so can substantially advance the field in both theoretical and analytical respect.

The present article primarily focuses on demonstrating the theoretical contributions to be gained by a more systematic use of adult development perspectives on corruption. Even though we also spell out some important practical implications of the approach proposed here, and provide a number of empirical examples, a more detailed discussion of how the MHC can be applied on different aspects of corruption research is beyond the limits of this paper.

The following article first briefly presents the Model of Hierarchical Complexity, which we propose as a theoretical lens to support a more complex, more integrative, and thus more effective corruption analysis. The remainder of the article is made up of two main substantial parts, devoted to a more detailed account and consideration of the implications of applying the MHC to corruption analysis, research and practice. The first of those sections focuses on how to analyze corruption itself with the MHC; the second section offers a meta-theoretical framework for re-evaluating scientific and political discourses about corruption. The paper closes with a general discussion of the findings and concluding remarks.

» PRESENTATION OF THE MODEL OF HIERARCHICAL COMPLEXITY

The Model of Hierarchical Complexity (MHC) is a model in mathematical psychology developed by Michael Commons since the 1980's (Commons, Trudeau, Stein, Richards, and Krause, 1998) on the basis of an earlier version, the General Stage Model (Commons and Richards, 1984a, 1984b). The following summary is adapted from the Introduction to the MHC in World Futures (Commons 2008). Permission to use the passages quoted here has been requested from the author. It is a quantitative behavioral developmental theory offering a standard method for examining universal patterns of evolution and development both in humans and in broader natural and social contexts. As a quantitative model, it accounts for the developmental content-based stage progressions evident in over 33 currently known structuralist developmental models by using more abstract terms and definitions than the former. Among the models included by the MHC are Jean Piaget's model of cognitive development, Robert Selman's model of the development of social perspective taking, Kohlberg's model of moral development and various models of self and identity development, to name only those which appear most important for understanding and analyzing corruption.

With those more domain-specific developmental models, the MHC shares the idea that complexity development can be defined as progression within a non-arbitrary hierarchical sequence of increasingly complex stages of performance by which the behaviors observed in the respective domains of development are ordered. This means that actions at a higher order of hierarchical complexity:

a) are themselves defined in terms of actions at the next lower or earlier order of hierarchical complexity (creating a hierarchy of increasingly complex actions that may be taken),

b) organize and transform the lower-order actions

c) produce new kinds of organizations out of lower-order actions in a non-arbitrary way.

Table 1. Stages of hierarchical complexity according to the MHC

	Order of stage	Possible operations and competences and their results
15	Cross-paradigmatic	Coordinates and crosses paradigms, builds new fields of knowledge (consisting of two or more paradigms)
14	Paradigmatic	Coordinates, integrates and synthesizes meta-systems (fields of knowledge), builds paradigms, requires high degree of decentration
13	Meta-systematic	Compares and coordinates various systems, builds meta-systems out of disparate systems, as well as meta-theories (theories about theories)
12	Systematic	Multiple relations between abstract variables, considers relationships in contexts (→ building systems)
11	Formal	Coordinates two abstract variables, calculates the influence of one variable on another one, solves problems with one unknown using algebra, 1-dimensional linear logic (if-then) and empiricism
10	Abstract	Builds abstract concepts and variables out of finite classes of concrete phenomena (time, place, act, actor, state, type), makes and quantifies propositions: logical quantification (quantifiers: all, none, some), categorical statements/stereotypes (e.G. "We all die")
9	Concrete	Full complex arithmetic (long division, short division), 2. Person perspective: takes and coordinates perspective of other and self, follows complex social rules, forms cliques, plans reasonable deals, conceives history and geography
8	Primary	Simple logical deduction and empirical rules involving time sequence, simple arithmetic (adds, subtracts, multiplies, divides, counts, proves), does series of tasks on its own
7	Pre-operational	Simple deductions, tells stories, counts events and objects up to 5, combines numbers and simple propositions, connects the dots, follows lists of sequential acts
6	Sentential	Chains words (coordinates words and names), imitates and acquires sentences and sequences; follows short sequential acts, pronounces numbers in correct order, acquires pronouns: subject (i), object (me), possessive adjective (my), possessive pronoun (mine), and reflexive (myself) for various persons (i, you, he, she, it, we, y'all, they)
5	Nominal	Uses words and names for things (coordinates and relates concepts), single words: exclamations, verbs, nouns, number names, letter names
4	Sensory-motor	Responds to stimuli in a class successfully and non-stochastically, forms simple concepts, morphemes (coordinates schemes)
3	Circular sensory-motor	Schemes (touch, grab, shake objects, circular babble, …), coordinates perceptions and movements, forms open-ended proper classes, phonemes, archiphonemes
2	Sensory or motor	Discriminates in a rote fashion, stimuli generalization, perceives and views objects or moves; moves limbs, lips, toes, eyes, elbows, head
1	Automatic	Engages in one action at a time. Cellular activities: sensing, effecting
0	Calculatory	Exact computation only, no generalization, human-made programs manipulate 0, 1; not 2 or 3

The next higher order actions thus cannot be accomplished by the respective lower-order actions alone. Rather, the higher-order action coordinates the actions of the next lower order by a mechanism with a higher degree of efficiency. "Thus, *hierarchical complexity* refers to the number of recursions that the coordinating actions must perform on a set of primary elements" (Commons, 2008) (see table below).

As a formal theory for scoring the complexity of behaviors, the MHC quantifies the orders of hierarchical complexity of tasks based on mathematical principles of how information is organized. The model's basic terms and categories have been well explained by Commons, Gane-McCalla, Barker & Li (2014) and do therefore not have to be outlined here (See Coombs, Dawes, and Tversky, 1970; Commons and Richards, 1984a, 1984b; Commons and Rodriguez, 1990, 1993; and Lindsay & Norman, 1977).

"So *tasks* are understood as the activity of organizing information. Each task's difficulty has an order of hierarchical complexity required to complete it correctly" (Commons, 2008). Quantal in nature, tasks are either performed correctly or not completed at all. There is no intermediate state or intermediate performance (even though there are transition steps).

"This objective, quantal feature of tasks and stages means that discrete ordinal scores can be assigned to them" (Commons, 2008). Hence, as a quantitative behavioral developmental theory,

the Model of Hierarchical Complexity includes a validated scoring system (see Dawson-Tunik, 2006, for hierarchical complexity validation studies).

Since hierarchical complexity applies to any event or occasion in which information is organized, the kinds of entities organizing information that can be studied by the MHC include humans and their biological systems as well as their social organizations, non-human organisms, and machines (for example computers). "The reason why it applies so broadly is that within its mathematical method of measuring tasks, scoring does not depend upon the content of the information (e.g., what is done, said, written, or analyzed) but upon how the information is organized". This makes the MHC fairly universally applicable cross-culturally and even cross-species, in any context. "Moreover, because the MHC's stages are conceptualized in terms of the hierarchical complexity of tasks rather than in terms of mental representations (as in Piaget's stages), it eliminates dependence on mentalistic, cultural, or other contextual explanations. Thus, the highest stage represents successful performances on the most hierarchically complex tasks rather than moral or intellectual maturity". However, as far as moral development is concerned, validation studies have shown a very strong relation between traditional scoring according to Kohlberg's Moral Judgment Interview (MJI) and the MHC's Hierarchical Complexity Scoring System (HCSS). Actually, the HCSS

is even more exact than the MJI, because it provides absolute scaling due to its content-independence (Commons, Galaz-Fontes & Morse, 2006).

The formal orders of the model are presented in Table 1. In order to analyze corruption and attitudes towards corruption with the MHC, tasks have to be defined which identify the concrete behavioral demands on each stage of performance. This will be done in the following sections.

» APPLYING THE MHC TO CORRUPTION — THEORETICAL AND ANALYTICAL GAINS

The following section will demonstrate how corruption can be understood in a more complex and thus, more adequate way by using the MHC. This will be done with regard to two aspects: First, in view of corrupt behavior itself, i.e. by looking at how behavior changes as action logics increase in complexity. Second, the model's descriptions of action logics also show how our understanding, perception and evaluation of corruption change as reasoning becomes more complex. Table 3 below illustrates how behaviors and their understanding increase in complexity on the different levels identified by the MHC. For the purpose of the present article, we have limited this discussion to the most common behaviors, i.e. those between primary operations (stage #8) and meta-systematic

Table 2. Correspondence of Stage Models (Kohlberg — MHC)

MHC stages			Kohlberg stages of moral development	
15	Cross-paradigmatic	(7)	(hypothetical)	
14	Paradigmatic	6	Universal ethical principles	Post-conventional morality
13	Meta-systematic	5	Social contract (may conflict with moral principles)	
12	Systematic	4	Authority and social-order maintaining, law and order	
11	Formal	3/4		Conventional morality
10	Abstract	3	Social expectations, interpersonal accord and conformity, good boy/girl	
9	Concrete	2/3		
8	Primary	2	Exchange, self-interest, what's in it for me?	Pre-conventional morality
7	Pre-operational	1/2		
6	Sentential	1	Obedience and punishment	
5	Nominal	0/1		
4	Sensory-motor	0		
3	Circular sensory-motor	-1/0		n.a
2	Sensory or motor	-1		
1	Automatic	-1/-2		
0	Calculatory	-		

Note. This table has been adapted from Commons & Sonnert 1994 and Tuladhar and Commons, 2014

operations (stage #13), since behaviors less complex than #8 and more complex than #13 are not yet relevant for analyzing corruption as of today. In order to explain the theory contained in table 3, we will first focus on corruption as a behavioral phenomenon, while the second subsection deals with our perception of corruption on different levels of complexity. However, both aspects are closely linked due to the logical relations of reasoning and action.

The complexity of corruption as measured by the MHC — theoretical and analytical gains

Before turning to the MHC, we wish to briefly recall Kohlberg's model in order to make clear to what extent the MHC is more differentiated and goes beyond Kohlberg's stage descriptions. The relation between stage definitions in Kohlberg and the MHC is demonstrated in table 2.

In his neo-Piagetian theory of political development, Stephen Chilton (1988) has illustrated the dilemma of corrupt behavior as seen by Kohlberg's theory with an example. In the example, a bureaucrat's professional ethics (loyalty to the rule of law, stage #4 morality) is challenged by various arguments, each of them coming from a different complexity level of morality: stage #1 threats, stage #2 bribes or stage #3 appeals to friendship. An effective bureaucracy as an abstract system of rules needs moral justification, i.e. bureaucrats functioning at Kohlberg's stage #4, typically arguing that observing rules and laws is an end in itself, because otherwise, social order and/or the functioning of the respective organization would break down. The following example presents the dilemma of a bureaucrat functioning on the basis of a stage # 3 morality, and whose professional ethics is challenged on that same level:

Client: Why don't you just set aside those requirements? After all, I am a friend and neighbour of yours!

Bureaucrat: If I did that I would disappoint my boss, who is counting on me to follow the rules.

Client: How can you put your boss ahead of me, your old friend and neighbour?

Bureaucrat: (no answer).

A bureaucrat making moral judgments on Kohlberg's stage # 4 would immediately recognize the inadequacy of this demand. Our stage #3 bureaucrat, however, "might dimly feel that the client's appeals in terms of friendship or personal ties are wrong, but stage 3 counter arguments give no clear support" (Chilton, 1988), since his personal ethics of service are equally stemming from a stage #3 personal loyalty, instead of from the authority of the rules or laws themselves. This also explainswhy functionaries often show little or no remorse with regard to their corrupt actions. Rationalization as explored by some of the authors reviewed by Fein & Weibler (2014), is a strategy which only appears on stages higher than Kohlberg's #3, because it requires an awareness of discrepancies between one's actions and the moral codes defined by the respective organization or context (see the paragraph on formal stage reasoning below). The absence of both remorse and rationalization in many contexts (indications of stage 8 or 9 action logics, see below) therefore explains why corruption is such a widespread and often unquestioned phenomenon there.

Moreover, Chilton's example not only illustrates a typical discrepancy between the institutional logic of a bureaucracy and the actual action logic of the empirical actor, showing that "unless the

Table 3. Stages of complexity of corruption and corruption communication

Order of Stage	General behavioral competences	Forms of corruption (if applying)	Examples (past and present)	Attitudes towards corruption	Reaction to corruption/anti-corruption action
14 Paradigmatic	Coordinates, integrates and synthesizes meta-systems, builds paradigms, requires high degree of decentration	???	Non currently known	Understands reciprocal nature of social relations in a broad sense (general good) → no simple answers	Makes shareholders take their positions themselves → organizes inter-shareholder dialogue (including real or perceived enemies) in order to co-construct acceptable shared sets of precepts and to make corrupt practices unnecessary
13 Meta-systematic	Compares and coordinates various systems, builds meta-systems out of disparate systems, formulates high level principles, as well as meta-theories (theories about theories)	Advanced capacities of perspective taking → corruption more and more unlikely; highly developed actors using their enlarged systemic view to take advantage of less developed actors.	As complexity increases, corruption turns more and more corporate or political, not individual anymore; price fixing, gaming the market on a higher level of complexity; i.e., one can really game the system in an enduring institutionalized way	Ability to take all stakeholders' positions → understands inherent logics and structural supports and constraints of corrupt behaviors and systems, as well as shortcomings of other systems to deal with it	recognize the range of structural changes necessary, far more than agreeing on precepts → works on flexible, stage adequate solutions to meet demands and enhance complexity development of actors on each stage. Solutions are beyond "one size fits all"
12 Systematic	Multiple relations among formal variables, considers relationships in contexts (→ building systems); more complex interpersonal relations, more abstract concepts (e.g. transparency, accountability, social justice, ethics etc.)	Understands complex systems in different domains (such as legal, market, financial system) → invent systems to skirt market and transparency rules	Speculation, monopolistic practices, price fixing, gaming the market Western businesses in developing countries, giving bribes because "this is the way the system works"	Corruption seen as counterproductive, dysfunctional and undermining economic systems, raising costs and prices, spirit > letter of law own values, experiences tend to be generalized and projected onto others	→ anti-corruption legislation, advanced transparency and accounting practices Effective competitive systems reduce corruption
11 Formal	Coordinates two abstract variables, calculates the influence of one variable on another one, solves problems with one unknown using algebra, 1-dimensional linear logic (if-then) and empiricism	Coordinates own interests and the risks of getting caught → ego-driven rational choice behavior serving own interests and trying "to get the most out of the system", strategic behavior to prevent getting caught.	Bureaucracies, including parastatals and other non-state actors, with or without free competition "beating the system" in the Soviet Union and other socialist systems	Sees how to advance own interests without breaking the rules and regulations, recognizes threats to one's interests posed by other actors → corrupt behavior becomes a problem → bad consciousness if caught, strategies of rationalization	→ legal criminalization, procedures to prevent/penalize corruption are created, but not implemented very effectively because systemic causation is not understood. In some cases little or no persecution, because this is, after all, how "business gets done" and how influence is deployed
10 Abstract	Builds abstract concepts and variables out of finite classes of concrete phenomena, makes and quantifies propositions: logical quantification, categorical statements/stereotypes	Social norms (e.g. to be honest and give fair service) and roles (e.g. boss, leader, subordinate, officer, teacher) are understood, but contradictions between roles and norms go unnoticed. Bribes ARE social norms in most abstract stage contexts/societies	Developing countries, Tsarist Russia (19th century)	Contradictions to norms are not considered as problematic if one is in the "out group", bribes are mostly not considered as "corrupt". → No relevant discourse about corruption, unless by external pressure	"Corruption" is not a legal category of illegitimate behavior yet → no anticorruption action.
9 Concrete	Full complex arithmetic, coordinating two perspectives, follows complex social rules, forms cliques, plans reasonable deals, conceives history and geography	Perspectives of ego and other are coordinated → deals are made which benefit both sides. No consideration of other perspectives, needs or rules.	Italian Mafia, early 20th century, other mafia-like structures	Deals are the normal way things get done. Power and money determine the outcome of deals. No abstract rules and evaluations.	Corruption might be seen as a deal gone bad
8 Primary	Simple logical deduction and empirical rules involving time sequence, simple arithmetic, does series of tasks on its own	Behavior performs single tasks sequentially, without coordinating or reflecting them as interrelated or as objects, e.g. "I give you something", "you give me something" without logical connection between the two	(No difference between "corrupt" and other behavior other than obtained consequences of breaking specific rules)	One takes what one wants if one can get away with it	Complains to teacher, parents and other authorities about other people's bad behavior

institution's structure is preserved by people at the appropriate stage, the institution will regress to less developed forms"(Chilton, 1988). It also makes clear that it is precisely the transition from Kohlberg's stage #3 to stage #4 action logics which is crucial for overcoming and preventing corrupt behaviour in organizations. However, Kohlberg's model does not offer categories for describing and analyzing the transition between those stages in more detail. This is where the MHC can make valuable contributions. From the general matrix of behavioral complexity presented in table 1, we can now deduce tasks defining the complexity of concrete behavioral demands and competences on each stage of performance.

The following overview of stage descriptions and action logics, as illustrated in Table 3 shall make clear that what we call "corruption," i.e. the misuse of public office for private gain,

» only comes into being as a social phenomenon after the abstract concepts of "public" and "private" have been formed (MHC stage 10), while the respective behaviors constitute the normal way of being and acting on MHC stages 9 and below;

» is only considered problematic or socially harmful after contradictions between social norms and individual behavior can be coordinated and dealt with in a non-arbitrary way (MHC stage 11) and

» that corruption can only be prevented or at least effectively reduced once efficient social systems (such as legal, financial, market systems etc.) are in place and functioning as the dominant social, political and economic structures, i.e. supported by a sufficiently large number of people (MHC 10). This is also where most discourse about corruption takes place.

» Moreover, adequate and sustainable solutions of the problem of corruption are likely to be reached only on the basis of at least meta-systematic structures of reasoning and performance (MHC stage 13 and higher) which are able to understand the inherent logics of corrupt behaviors and to design stage-sensitive solutions beyond "one size fits all."

At this point, we have to stress that as a rule, organizations, like societies in general, "are comprised of individuals operating at multiple stages of development in various domains" (Ross & Commons, 2008). Thus, organizations, as well as "political cultures and social systems display concurrent operations of several different stages. There are many overlapping systems and relationships among different people and entities. That fact has understandably contributed to analytic and policy confusion" (Ross & Commons, 2008). At the same time, there are always modal stages, i.e. stages at which most individuals operate within organizations, societies and governments and which thereby characterize the stage at which the respective entities are likely to operate as a whole (Commons & Goodheart, 2007).

We will now explain the behaviors and their underlying action logics on each stage of complexity as contained in Table 3 in more detail and discuss their implications in view of corruption, as well as corruption control. The following stage descriptions are based on and in part quoted and/or adapted from Ross & Commons (2008).

Concrete behavior—corruption in organizations "avant la lettre."
Concrete stage 9 reasoning "focuses on events, people, and places that are personally known" (Ross & Commons, 2008, p. 484). Individuals, societies and (members of) organizations functioning at this stage are preoccupied "by subsistence concerns and demonstrate short time horizons.Social behavior is [therefore] characterized by reciprocal exchanges involving concrete goods and services, and simple social rules. Dyadic relationships are prevalent (e.g., to plan deals, trade favors, and barter). Others' perspectives are considered only if those others affect oneself or one's close group or enable deals that both parties regard as fair" (Ross & Commons, 2008, p. 484). Since there are no abstract concepts such as public/private, a greater common good or rules about right or wrong yet, societies or organizations functioning at a (hypothetically) pure concrete stage do not have bureaucracies, administrative structures and civil services in a Weberian sense.

This means that individual actors reasoning at this stage cannot take the perspective of the organization as an abstract whole or consider its overall rules and codes of conduct as long as the latter are not identical with their own immediate needs and interests. Rather, their thinking and behavior is organized around we-groups, the members of which are personally close (like my family, my tribe, my clan, my friends etc.). Therefore, if individuals at the concrete stage hold "public" offices, they tend to treat them as personal property, or as a means to expand their personal property. In fact, this was the dominant attitude towards public offices in all pre-modern societies all over Europe (Schattenberg, 2008). It is therefore not surprising that this action logic can still be found in developing countries all over the world.

As to organizational logics on the concrete stage, Ross & Commons (2008, p. 484) state that "at some point, some approach to formal government is introduced" there. But its purpose or function is not to implement abstract principles or organizational goals, "but rather the power and wealth of its leaders, and only to some degree the protection of its subjects. At this stage, specific officials (e.g., a king, leader, warlord, president, or minister) essentially 'are' the government from the concrete stage perspective. This is because roles are not separated from the ones who fill the roles. They do not have to be, because leaders are personally known or known of, and followership is based on personal and economic ties, not roles. Without concepts of contracts or title to goods, government is not needed to regulate transactions; physical possession constitutes ownership and power. Despite possible appearances of a form of central government, rule is exercised in traditional ways: making deals and exerting raw power in the 'friend or foe' mode" where the outcome of deals is determined by power and money (Ross & Commons, 2008, pp. 484-5), as it is mostly the case, for example, in Mafia groups all over the globe.

While higher stage individuals, governments and international bodies commonly judge concrete stage societies' efforts to have and run organizations and governments as corrupt, from their own perspective deals, "bribes and 'under the table' reciprocal arrangements are the normal way to conduct affairs" (Ross & Commons, 2008, p. 485). Since neither roles nor abstract concepts such as "private/public", nor formal rules exist on this (hypothetical) stage of individual and organizational complexity development,

they cannot conflict with one another. Thus, the term "corruption" does not exist as an analytic or (dis-)qualifying concept. In our view, it is important to understand this, since in the theoretical perspective of complexity development, there is no point in trying to convince concrete stage individuals to conform to abstract rules or ethics codes whatsoever. Rather, strong personal power-based authority is the strategy that works best with them in order to eventually achieve norm conformism.

In Kohlberg's model, the concrete behavioral logic corresponds to the transition between stages 2 and 3 of moral judgment. While stage 2 morality primarily asks "what's in it for me", and perceives others in relation to what they can contribute to the satisfaction of one's own needs and interests, stage 3 morality is defined by conformism to reciprocal expectations and good interpersonal relations. A good example for a concrete stage culture is the "ethics of southern Italian mafia of the early 20th century as described by Pino Arlacchi (1989; see also Paul & Schwalb, 2011).

Abstract stage behavior and corruption in organizations. It is only with the movement to the abstract stage and the development of social norms that the use of force is seen as illegitimate and that practices such as bribes may be considered corrupt – if they go against the respective social norms in place, which, however, is not always the case. At this level of complexity, abstract thought develops as a new competence, forming "variables out of finite classes" of concrete phenomena, and making "quantifying abstract propositions" (Ross & Commons, 2008, p. 485), classifications and generalizations. Only on this basis can abstract ideas and concepts such as public/private, as well as social roles and rules be understood. "People performing at the abstract stage value social norms". This is why they begin to form bureaucratic organizations Very often, they are quite attached to religious values and ideas like personal honour. This means that they can now "negotiate by *trading normative values* (unlike Concrete stage 9's dealing in tangible currencies from money to animals to people)" (Ross & Commons, 2008, p. 486).

However, those abstract roles, rules and values cannot yet be coordinated with one another in a non-arbitrary way. This is because in abstract thought as understood here, "individual rules can be conceived to accomplish a desired end, but the method to implement the rule cannot be conceived. (…) A rule can thus be explained and followed, yet contradictions with other rules or norms go unnoticed" (Ross & Commons, 2008, pp. 485-6). This may lead to dysfunctional behavior in the sense that different norms are played off against each other. "For example, a bureaucrat may be as faithful to the norm of charging bribes (because that is the way things get done), as to the rule to be honest and give constituents fair and equal service" (Ross & Commons, 2008, p. 486).

This is especially the case since abstract stage "loyalties are [often] unquestioned" and based on "group memberships which help people form their identity at this stage" (Ross & Commons, 2008, p. 485). While in contrast to concrete stage 9 we-groups, abstract stage 10 group associations begin to take the shape of membership in social relationships with others that are now possible also without physical contact to other members (for example in political parties, trade associations and unions, and religious organizations), a clear method or principle for deciding between

conflicting loyalties is still missing. Loyalties are rather based on groups' or leaders' belief systems or ideologies, often connected to dualistic assertions, prejudices, stereotypes about, and definitions of the "in-group" and the "out-group. Therefore, "strong, paternal-type leaders, often charismatic, tend to be preferred, on the assumption that they will take care of their children/followers and keep the group or society harmonious and fair". In cases "when real differences cannot be solved any other way, abstract stage negotiations can also agree to live with them to preserve harmony" (Ross & Commons, 2008, p. 486). This is why abstract stage organizations and societies often appear to be inconsistent and/or chaotic to outsiders, especially to observers on higher stages of reasoning complexity.

This action logic probably describes most of the organizational practices in pre-modern and/or developing societies, both in European history and in the present, and has only gradually been combined with or replaced by more complex logics (Fein, 2012; Schattenberg, 2008; Fleck & Kuzmics, 1985). Moreover, it can be assumed to play a considerable role in organizations still today. In Kohlberg's terms, abstract reasoning roughly corresponds to stage #3 conventional morality and is thus, as explained in the beginning of this section, not adequate in view of preventing corruption in organizations. For even though bribing starts at and is most typical of the concrete stage, it also exists at the abstract stage, where its power and influence depend upon the culture as to what is socially normative. In some cultures, bribes clearly *are* the social norm. (Note that even though the practice of bribing is lessening at higher levels of complexity, it does extend up even into the systematic stage. But in cultures in which bribes are not tolerated, then at the abstract stage, they are not tolerated by downward assimilation, i.e. because a sufficiently strong formal (stage 10) action logic and value system exist in the respective organization or social context which has enough authority to make sure that bribing is criminalized.

So in view of implementing ethics codes vis-à-vis stage 10 individuals inside organizations, it is crucial, first, to analyze where and to what extent abstract reasoning is present in the respective organization, and to what degree the actual norms of the abstract stage reasoners are either compatible with or contradict the desired codes of conduct. Second, it might be advisable to appeal to the abstract stage's sense of loyalty by making clear that the desired code of conduct is an essential part of organizational identity and vital for its survival, well-being and inner harmony.

Formal stage behavior and the criminalization of corruption in organizations. The first action logic clearly discriminating and criminalizing corruption as an inefficient behavior is the formal stage #11 reasoning. It is characterized by the ability to coordinate two abstract variables in a non-arbitrary way and therefore, to see and to avoid contradictions between behavioral norms and rules, as well as between those rules and actual behavior. Formal stage reasoning is more complex than abstract reasoning, "because it involves solving problems by using logic, mathematics, and empirical investigation in order to find out what is true. What is considered true [here] is thus based on forming relations out of variables, where logic is linear and one-dimensional, because only one input variable can be considered at one time" (Ross

& Commons, 2008, p. 486). For example, people reasoning at this stage "prefer uncorrupt practice once they deduce that they can save money and have more predictability" this way (Ross & Commons, 2008, p. 494).

These new logical relations, along with enlarged capacities of social perspective-taking, help people to distinguish between social roles and the individuals who hold them, and to understand "logical cause-and-effect-based regulations and procedures", as well as their benefits and consequences. "Roles and procedures thus come to be viewed as logical necessities for organizations and government to function well enough to succeed" (Ross & Commons, 2008, p. 494). People thus gradually learn to communicate through more impersonal contacts and to rely on formal institutions in their everyday affairs. In result, abstract stage corruption is reduced, because once formal regulations define power in legal terms, supported by systems of checks and balances, "the ability of individuals to exercise personal power over public resources declines". Moreover, "a key government task in the change from abstract to formal stage is to legislate a social contract that takes over the functions of the previously ubiquitous informal systems" (Ross & Commons, 2008, p. 494).

In stage 11 organizations or societies, formal economics and laws are also advanced by the formal action logic's "empirical interest in increasing productivity, training, and wealth distribution". Moreover, it discovers "that the existence and enforcement of criminal and civil law promotes trade and investment. This connection is made easily at this stage because each is a simple empirical relationship between two abstract variables" (Ross & Commons, 2008, p. 486). Members of formal stage societies therefore begin to explicitly demand the rule of law to prevent (…) corruption" and to increase efficiency of public institutions). Ross & Commons (2008) add that "depending on the culture, it takes time and courage for citizens to publicly voice such demands. Such behavior may risk one's status in the patronage systems people have long relied on" (p. 494). So if the formal action logic represents the dominant culture of a society or inside an organization, it has the capability to bring about and maintain efficient impersonal structures, systems and bureaucracies for regulating social and organizational life. "Extensive written laws and regulations" tend to be "implemented in 'letter of the law' fashion" (Ross & Commons, 2008, p. 486). To the extent that rules and laws become "effective in moderating organizational crime", including corruption, competition turns more "civil", and it is more and more "the contingencies of the marketplace which control social relations and status". Therefore, formal stage reasoning is crucial for establishing and maintaining public infrastructures. For the same reason, "this stage is the objective of many efforts to introduce market economy and democracy."

However, when formal stage regulatory ideas are exported to non-Western countries", to contexts that have not known them before or that have not developed them by themselves, "there may be too few persons performing at the formal stage to understand how procedures are supposed to work or the underlying logic (e.g., separation of legal powers or administrative duties)" (Ross & Commons, 2008, p. 486). If the target context is mistaken for a formal stage context, "the new forms of government or business procedure may just provide new facades to which conventional

behaviors of patronage adapt and persist, usually even more effectively because access to new resources is available. Ross & Commons, 2008 argue: "For example, the formal concept of employees on payroll is used to pass resources to clients, often as 'ghost employees' who do not work for the employer. (…) Bureaucracies become engorged through such arrangements. Because in-group ties are stronger than other ties in abstract settings where formal stage structures are imported, many people are often less successful at distinguishing an employment role from a political party role, for example, party loyalty trumps formal role responsibility" (p. 486).In socialist systems, we witnessed a curious mixture of formal bureaucratic systems and partly less-than-formal cultures and action logics which is why those systems often did not function effectively (see Merl, 2012, 2010, 2008; Voslensky 1984).

On the other hand, there is also a formal stage corruption proper, for "people who use formal reasoning are good at using rules to find or create loopholes to implement their own strategies" (Ross & Commons, 2008, p. 486). In this case, one's own interests are coordinated with another variable such as the risk of getting caught. At the same time, formal reasoners are not very good at anticipating the consequences of their actions, because a more complex systemic perspective is not yet developed. Due to missing systematic coordination of variables, formal stage actors "may be clever at 'cooking the books' to hide bribes, yet not foresee how they will either still get caught" (Ross & Commons, 2008, p. 486), or how their behavior (further) undermines the working of the system as a whole. In both cases, rationalizations are a typical formal stage strategy to justify one's behavior or to "buy oneself out of trouble."

In this sense, the formal action logic is an important step towards Kohlberg's stage 4 morality, but not yet this morality itself, because it does not yet see and take into account the more general systemic consequences of one's own behavior as the former would do in a rather strict manner.

Systematic stage behavior, corruption and corruption control in organizations. The competence to simultaneously coordinate multiple variables only appears on stage 12, which is characterized by systemic reasoning and acting in more complex contexts, social relations and time horizons. Actions at the systematic stage 12 (and within action logic research, we also consider thinking as an action) are defined by the "coordination of more than one variable as input and the consideration of simple relationships in context. These coordinations and considerations construct multivariate systems, matrices, and webs of causation, resulting in more complex societies" (Ross & Commons, 2008, p. 487), as well as more complex theories about these societies.

"In systematic stage societies, systems of formal relations are coordinated among the legal, societal, corporate, economic, scientific and national spheres. Because at this stage, organizational systems are complex enough to address and achieve multiple goals simultaneously, society is predominately lawful, and advanced accounting practices make business relatively transparent. Markets, stock exchanges, and the like produce complex impersonal relationships among people, and more intricate laws and regulations stabilize markets and attempt to prevent monopolies" (Ross & Commons, 2008, p. 487). So at this stage, corruption is

generally further reduced, yet still existent, and becomes more sophisticated. Moreover, the fact that people reasoning at the systematic stage expand their perspective on themselves and the society still further has several consequences in view of corruption and corruption control.

Since "people can now consider a multivariate combination of such factors as the rule of law, fear of exposure, preservation of image, methods of reporting, and market pressure" (Ross & Commons, 2008, p. 487), behaviour becomes more conscious and more differentiated, taking into account broader social horizons, interests and constraints. This leads, for example, to the introduction of professional norms which define being a professional as having a role independent of personal affiliations and conflicts of interest. Also, "more highly abstract concepts [are formed], such as transparency, accountability, social justice, and sustainability". On this basis, systematic reasoning "can conceive systems of transparency and control to reduce corrupt practices" (Ross & Commons, 2008, p. 487). In fact, the systematic perspective is the first to recognize corruption as a systemic problem that has to be fought, because it is understood as being counter-productive and dysfunctional for the working of the system as a whole. This is why only stage 12 systematic reasoning fully corresponds to the definition of Kohlberg's stage 4 morality, which we identified earlier as a necessary condition for overcoming and preventing corruption in organizations. As Kohlberg's stage 4 morality, systematic stage anti-corruption discourses would argue that corruption is detrimental to the functioning of organizational systems as well as to the broader social systems in which the former are embedded, and that this is why it has to be addressed in a principal and consequent way. As a result of dysfunctional elements being identified, institutions start to function better once the systematic action logic has become the dominant culture. Also, substantial criteria gain importance over formal ones. For example, "applications of laws are now more 'in the spirit of' than 'the letter of' the law" and procedures, and governmental processes become more "orderly and fair"" (Ross & Commons, 2008, p. 487). Democracy is therefore valued not only for its formal and efficient rules as on the previous level of reasoning, but for just outcomes.

However, the systematic stage's enlarged perspective also entails two kinds of ambivalent consequences. First, systemic reasoning can not only promote corruption control, but also "conceive systems to skirt efforts to enforce transparency" (Ross & Commons, 2008, p. 487). Frequent types of stage 12 corruption are speculation, monopolistic practices, price fixing, and gaming the market, as well as, on the international level, multinational corporations using bribes to "get business done" in non-Western countries, often rationalizing that "this is the way the system works", as was the case with Walmart in Mexico, for example. This is why "this stage can neither succeed in entirely escaping transparency measures nor eliminate efforts to sabotage attempts to institutionalize transparent practices and reduce corruption" (Ross & Commons, 2008, p. 487).

Second, due to missing meta-systematic competences to coordinate several systems, i.e. to take a self-critical, distant look on their own system of values and institutions, "legislators, judges, and administrators" at the systematic stage tend to project their own ideals of organization and their experiences of government onto others and other contexts "in a logical, but non-empirical or scientific manner" (Ross & Commons, 2008, p. 487). In a related way, systematic reasoners also tend to "assume a common value system" across societies "or, where values differ, that their own, i.e. the value system" of the respective researcher," international body, legislator, or government official is 'right'", i.e. most progressive, "and that of the others is 'wrong'" (Ross & Commons, 2008, p. 487). This tendency often motivates the export of systematic stage (sometimes declared "Western") systems to non-Western or earlier-stage settings, "where they fail… because they are systematic stage" ideals, reasonings and action logics which "are imposed on and expected from earlier-stage settings to no avail" (Ross & Commons, 2008, p. 487).

A more adequate and thus more efficient strategy would thus be to take stage differences between reasoning and action logics into account both in theoretical and in corruption control practice. This, however, requires an even more complex and decentered perspective, which is only possible at the meta-systematic stage.

Meta-systematic stage reasoning and behavior and its handling of corruption in organizations. Following the MHC's conceptual logic, the meta-systematic stage 13 is defined by actions that "compare systems", amongst them systematic stage perspectives, and "create supersystems out of systems of relationships" (Ross & Commons, 2008, p. 487). This means that meta-systematic reasoning not only sees the limitations of the systematic action logic, but it also recognizes and considers the other action logics as *systems* of thought and action in their own right, following their own internal logics and dynamics and each having its own merits and shortcomings. Because it is even more self-reflexive and more detached, meta-systematic reasoning no longer sees its own value system or reasoning habits as a universally desirable (or realistic) "fit" for each context and every organization. Instead, it is able to take the perspectives of the action logics empirically functioning in each context and, accordingly, to handle the internal logics and dynamics of various types of corrupt behavior appropriately. In result, meta-systematic reasoning is more likely to generate solutions that are more sustainable, because they can flexibly match the specific problems on each stage and in each context.

In this respect, a central challenge consists in meeting the demands and motivational structures of the different logics of reasoning and action. This means, for example that concrete reasoners can most likely be convinced by power, and abstract ones by authority, whereas formal action logics have to be motivated by reasonable self-interest, and systematic reasoners by appealing to their sense of professionalism and social responsibility. Only on the basis of taking into account this complexity dimension of reasoning and behavior (action logics), we claim, can organizational ethics programs and codes of conduct be implemented successfully and sustainably. And only on these grounds can further complexity – and thus, ethical development of the respective actors eventually be enhanced.

Even though similar political and organizational meta-systems are empirically rare and, where they exist, still imperfect and fragmentary, they attempt to "incorporate the much higher amount of complexity involved in adequately qualifying any system of [thought, action and] duties, (…) beyond 'one size fits

all'". Moreover, only the meta-systematic perspective proposed here fully recognizes that the latter kind of efforts is "limited by assumptions that do not stand up to the order of complexity which actually must be addressed, and are therefore condemned to fail" (Ross & Commons, 2008, p. 488).

Hence, our analysis suggests that meta-systematic stage perspectives are a necessary and extremely efficient tool for more adequately understanding both empirical phenomena of corruption in organizations and the way theorists and practitioners have tried to cope with those phenomena. Not only can governments and organizations do a more complex job based on a meta-systematic action logic, able to handle behavioral, psychological legal and institutional systems simultaneously. Meta-systematic theory building as proposed here also provides insights into the inherent patterns of structural complexity of current discursive and scientific constructions of corruption. It can therefore help to re-evaluate and more adequately assess the plurality of attitudes towards corruption and to overcome disciplinary reductionisms. This aspect shall be discussed in some more detail in the following subsection.

Analyzing the complexity of attitudes towards corruption: public and scientific discourses on corruption as measured by the MHC — theoretical and analytic gains.

The previous section has made clear that "corrupt" phenomena only appear at a particular level of complexity of reasoning and behavior, that they change in character while action logics become more complex and more differentiated, and that they increasingly become objects of reflection in result of this process of complexity development. The following section summarizes the essence of the stage descriptions given in the previous sub-section, placing a special focus on the complexity of *perceptions of* and *perspectives on* corruption on different stages. On this basis, we then discuss how some of the contributions to scientific discourse on corruption reviewed elsewhere (Fein & Weibler, this issue) can be re-evaluated by using the MHC. However, we wish to stress that the examples from literature on corruption cited in the following section are merely intended to illustrate the respective structures of perceiving and theorizing about corruption presented here. Note that we do not claim to thereby give comprehensive evaluations of the publications cited with regard to their overall structural complexity. The latter would demand a much more systematic analysis of the respective contributions which is beyond the scope and the purpose of this article.

In view of analyzing the complexity of attitudes towards corruption, a general hypothesis is that perceptions and understandings of corruption become more comprehensive, more encompassing (more "holistic"), and thus more adequate as perspectives become wider, i.e. as more aspects of the empirical phenomena are included, and as social actors' perspectives are explored in more depth within a synergetic research design. Besides the number and kind of factors and variables considered, further criteria for measuring the complexity of discourse are its degrees of self-reflexivity (called "subject-object balance" by Robert Kegan, 1982) and contextualization, the nature of anthropological statements or assumptions that are, and the relation between analysis and evaluation of corrupt phenomena. So let us briefly go through the stages of corruption as defined by the MHC in table 3 again and look at how each of them meets and treats those criteria.

Concrete stage 9 and earlier perspectives clearly have no discourse on corruption. This is because what higher stage reasoning perceives and considers as "corrupt" is the normal way to get things done in those action logics. Since the dominant perspective is egocentric here, there is no self-reflexivity on concrete or earlier stages, nor is logical scientific reasoning possible. Also, normative evaluations do not occur, since abstract norms and concepts (such as right/wrong, corrupt/moral) have not been formed on these stages.

Abstract stage 10 reasoning can conceive abstract norms and rules, such as fairness and honesty. It can thus also distinguish "corrupt" from "moral" or "correct" behavior. But since those concepts cannot yet be coordinated in a non-arbitrary way, contradictions to and between those norms and concepts tend to go unnoticed. In result, neither those contradictions nor the respective behaviors are perceived as problematic. In most abstract contexts, bribes are therefore not considered as "corrupt". Reflexivity is still limited on this stage and does not include rational/logical scientific reasoning. As a consequence, there is no relevant discourse about corruption on this stage, unless by external pressure (downward assimilation). Instead, abstract stage reasoning, on the one hand, tends to make categorical assertions in the sense of stereotypes without empirical basis or logical/theoretical reflection (e.g. "all evil people go to hell"). On the other hand, it clearly treats other problems as more important than corruption, such as personal honor and being on the right side of the bar. For evidence from Russia see Fein (2012).

Formal stage 11 reasoning therefore has to be considered as a major breakthrough in our collective way of sense-making and understanding the world. Based on what Piaget identified as formal operational thought, this structure of reasoning is able to coordinate two abstract variables in a non-arbitrary way. This means that formal reasoning is capable of and interested in determining the relations between variables in a scientific, i.e. logical, reliable, and often empirical way. This competence is the very basis of and an essential component and condition of scientific practice until today.

With regard to the perception and analysis of corruption, this has two implications. First, its capacity to see and coordinate the differences between norms, between behavior and norms, as well as between personal interests and public and/or organizational goals makes it the first reasoning structure able to clearly define and identify particular phenomena as "corrupt" both on a theoretical and practical level. Hence, corrupt behavior becomes conceived of as a problem here. As a consequence of the formal competence to make clear evaluations, corrupt behavior tends to be legally criminalized as conflicting with and as inefficient compared to formal rules and procedures. Moreover, such rules and procedures are introduced to regulate how violations shall be dealt with. However, in societies that acquire formal level thinking for the first time, legal criminalization mostly remains formal (i.e. letter-of-the-law) and is not yet accompanied by systematic practices to ensure its practical implementation (spirit-of-the-law).

At the same time, corresponding discourses do arise both on the social/public/political level and in the scientific domain,

discussing, not only how corrupt behavior can be punished, but also how it can be explained. In this respect, formal reasoning may identify specific factors as particularly relevant, for example factors situated in the personality of the respective perpetrator. It will then discuss how those factors can be manipulated in order to fight corruption. In fact, this is what a large portion of research on corruption continues to do: It tries to explain corrupt behavior in relation to or as a result of the interaction of specific, often quantifiable variables, for example the amount of wages paid, the size of bribes, the degree of a person's love of money, self-interest or behavioral control (see the examples cited in section above).

However, since formal reasoning can consider only one input variable at one time, its logic is linear and one-dimensional. This means that explanations on this level of complexity only consider causal relations between two variables at a time. As a result, they construct logical, yet rather unidirectional cause-and-effect relations which, from more complex levels of reasoning, appear to be reductionist, because a systematic perspective of the institutional contexts of corruption is still missing. This is also why formal legal criminalization is often not accompanied by systematic persecution and effective anti-corruption action on a societal level due to the lack of well-functioning state bodies able to implement the law. The Soviet Union and other developing countries are classic examples of existing formal anti-corruption legislation lacking systematic implementation.

From these perspectives, critiques like those voiced above may be raised, namely that corruption is studied by constructing plain variables out of complex phenomena that not enough aspects of the problem are considered, and that linear assumptions of causality between those variables are made. Another critique might be that discourse on corruption on a (hypothetically pure) formal stage does not explore the deeper, more complex dimensions of behavior and social relations and rather uses some form of rational choice theory to make simplistic assumptions about both of them, which are not empirically supported. Similar practices have been identified in a considerable number of publications on corruption in organizations as discussed above and by Fein & Weibler (this issue). As reasoning becomes more complex, those shortcomings are increasingly reflected and can thus be gradually overcome.

Systematic stage 12 reasoning takes the perspective of the system which it is part of, i.e. an organization, institution or society as a whole. This means that it sees and discusses corruption by asking what role it plays in the functioning of the respective system or organization. From that perspective, corruption is easily recognized as counterproductive, dysfunctional and harmful to systems based on the rule of law such as market economies and liberal democracies, since it undermines their rules, exchange mechanisms and codes of conduct.

An interesting exception, or rather variation of this kind of systemic stage reasoning is the analysis of corruption in the Soviet Union as proposed by sociologists like Stefan Merl (2008, 2010, 2012) and Michail Vozlensky (1987). It reveals the relationship between societal and organizational cultures and individual behavior. In socialist systems too, corruption was generally an illegal, yet widespread practice. "Corrupt" behavior (at the time called "blat" in Russia & Ledeneva, 1998, 2006) was often legitimized by social

actors by arguing that there was no other way to get things done, since the formal institutions did not work properly in the sense that they were not able to guarantee functional distribution of goods and services and to provide economic welfare. While social actors from the Soviet context themselves tend to ex post legitimize their own behavior by typical formal stage rationalizations (e.g. "in my case, it was not *blat*, it was help/friendship", "a favor is not illegal", "between friends the requests can be unlimited. […] If my best friend asked me something, I felt morally obliged and, in fact, preferred to compromise with my formal duties rather than break our relationship" etc., Ledeneva, 1998), researchers have argued that without corrupt or other illegal/"unethical" practices, the economy would not have worked at all. In other words, they declare those practices to have been functional elements of the socialist system in place. Of course, from a systematic perspective, the opposite (classic Kohlberg stage # 4) interpretation would equally be possible: If everyone had observed the rules, standards and political requirements (which in fact were indeed partly unrealistic), there would have been no problem, neither with corruption, nor with economic supply. Unfortunately, a more thorough empirical support of either position is impossible in this case.

However, in general, systematic stage discourse about corruption is connected with a clear negative evaluation. Research based on this reasoning logic therefore tries to combat corruption on a structural level, i.e. not only by punishing corrupt actors, but by putting into place complex systems of corruption control, for example through advancing sophisticated transparency rules and accounting practices. In other words, it tries to improve the (efficiency of the) system in any possible way by inventing new mechanisms of control and/or prevention or, inversely, by introducing incentives for ethical behavior. This is done on the basis of a more complex understanding of the interrelations between multiple variables determining individual behavior and the functioning of institutions.

Thus, behavior is now either regarded as *one* variable amongst others within a complex system of interrelations, or as a result of multivariate influences itself, and thus assumed to follow more complex logics. However, those logics are usually still studied on the basis of theoretical assumptions about behavior such as the ones made by rational choice or other behavioral theories, whereas the internal dimensions of behavior tend not be explored empirically in a more systematic way, for example in view of uncovering its inherent dynamics. This is because, in contrast to meta-systematic stage 13 reasoning, systematic stage thinking does not recognize behavioral logics as systems, and because multiple systems (action logics) are not yet coordinated with one another in a non-arbitrary way.

In connection with this stage's systems view and its sensitivity for multivariate interrelations, the notion of context is newly acquired. Since this concept has become an important element of the currently dominant systematic stage scientific culture and discourse, it is no surprise that the importance of context is stressed in many of the publications reviewed here – even if they do not consequently practice contextualization themselves. Note that indeed, systematic stage reasoning is not able to contextualize *itself* and its own functioning due to missing meta-systematic

competences. This is why its evaluations are often framed in moral terms or as appeals to ethical principles such as fairness and social responsibility. In fact, since on this stage, the spirit of the law is more important than its letter, more attention is put on establishing cultures that support the principles that are now "generally accepted" as ethical. This is why ethics codes defining rules of conduct and organizational best practices are very attractive to this kind of reasoning as normative guidelines for regulating and evaluating the behavior of organizations and their members. And because systematic reasoning does not yet contextualize itself, it also tends to generalize its own insights, perceptions, values and experiences once they have been found to be scientifically true, economically successful or ethically most progressive in a process of either rational or moral discussion, or economic competition, and tends to project them onto other actors and contexts. Often enough, this includes developing strategies to export systematic stage values and institutions, amongst others its systems of corruption control, to other parts of the world.

As might already have become clear by now, most of the current public and scientific discourse on corruption in western societies is functioning on this level of reasoning. Implicitly or explicitly acting on the assumption that modern capitalist economic systems and democratic political and organizational cultures are the most progressive and sophisticated forms of organization and governance, the respective systems and their internal mechanisms are analyzed in view of their performance, merits and shortcomings in various domains, but they are not contextualized themselves. This is why, for example, the stage dependent character of their concepts and evaluations remains unquestioned, and the structural nature of different kinds of "corruption" unnoticed. So while stage 12 typically does see various forms of corruption or corruption control as distinct, it merely treats them as types rather than as independent systems of thinking and acting (action logics) of their own. An example for this is Donald Lange's (2009) brilliant model identifying four types of corruption control, each "serving different functions" (autonomy reduction, reward and punishment, legal compliance and social conformity, and intrinsic motivation). While Lange does mention a connection with "individual differences […] such as differences in the stage of cognitive moral development", he does not systematically relate the workings of those types of corruption control (which exactly correspond to Kohlberg's stages 1-4) with the respective action logics that either bring them into existence or make them functional and efficient in particular contexts or with particular people, i.e. people functioning according to the respective action logics. And even though he rightly recognizes that "any particular corruption control type entails implicit assumptions about human nature" (2009), he does not give the bigger picture according to which all of those types of behavior are part of the same (more complex understanding of) "human nature" which can – and at the same time has to develop those structures one after the other. Similar comments apply to Seraphim Voliotis' study on the abuse of authority (2011), which offers an interesting typology of corrupt behaviors, yet again without integrating them into a coherent meta-system according to some overall, non-arbitrary principle.

The new, more complex kind of task that can be completed at the meta-systematic stage 13 is the capacity of building meta-systems out of systems, i.e. of coordinating different systems (among them scientific theories) with each other in a non-arbitrary way. In result, this order of reasoning is able to build meta-theories for organizing previously disparate theories in a way that makes visible the merits and shortcomings of each theory based on an evaluation of their respective structural complexity. In order to perform this task adequately, an even greater detachment and (self-) reflexivity is necessary. Moreover, this reasoning structure starts to practice second order contextualization both in view of analyzing "corrupt" (and other) behaviors and with regard to the ways those behaviors are dealt with by other logics of reasoning (theories about and attitudes towards corruption). In other words, meta-systematic reasoning contextualizes systematic stage reasoning itself. On these grounds, it recognizes how all theorizing and behavior is a function of its own internal structural complexity. It is thus decentered enough to take (and change between) different theoretical, as well as stakeholders' positions, i.e. to see and appreciate all of them both from the inside (according to their own internal logics instead of evaluating them on the basis of some external set of values) and at the same time to analyze and interpret them from a detached structuralist view.

In this sense, the theory and arguments presented here are meta-systematic in that they propose a meta-theory integrating various theories in a non-arbitrary way. The Model of Hierarchical Complexity provides tools for assessing the structural complexity of reasoning and behavior. It thereby helps to identify the possibilities and the limits of different behavioral logics by exploring the structuring mechanisms and internal logics which at the same time motivate and constrain empirical behaviors. Assuming that all behavior is consistent from its own point of view, meta-systematic reasoning tries to detect the principles, criteria and cognitive frames within which the respective behaviors can be seen and understood as "normal", logical and coherent. On this basis, the latter can thus be perceived as independent systems in their own right.

In view of dealing with corruption, meta-systematic understanding of the internal logics, supports and constraints of various forms of "corrupt" behaviors leads to a kind of societal macro-morality. This means, above all, a stage (or complexity) sensitive way of perceiving and reacting to them, thereby overcoming the structural shortcomings of other theoretical and practical systems trying to deal with those problems. Meta-systematic ethics or morality could thus be conceived of as a morality of adequacy, working on flexible, stage adequate solutions, in order to meet demands and enhance complexity development of actors, organizations and institutions on each stage. Hence, "societal macro-morality" also means that solutions generated on this stage of reasoning are beyond "one size fits all."

At the same time, due to its advanced capacity of perspective taking, to compare and coordinate various systems, and its ability to contextualize its own perspective, meta-systematic reasoning is also aware of the fact, that there are even more complex ways of theorizing about and dealing with corruption ahead, which have not yet been developed on a cultural level to a more relevant extent (and which are therefore left out of the discussion here).

So if we try to give an overall evaluation of the publications on corruption reviewed here and in Fein & Weibler (2014) from a meta-systematic perspective (more thorough individual evaluations are beyond the scope of this article), we can conclude that most of them are based on structures of reasoning situated between formal and systematic logics, with a few exceptions reaching into meta-systematic thinking. As a matter of fact, we often find mixed forms. This might be due to multiple interrelations between the respective researchers' own habits of reasoning and influences of particular scientific or disciplinary cultures which authors are part of. It may also be due to different researchers/authors working together using different orders of complexity according to which they understand the subject in question, so that their collaborated paper is mixed with orders of complexity. This hypothesis and/or experience was proposed by Sara Ross (personal communication, 2012, June).

In conclusion, in order to make clear that meta-systematic or similarly complex research perspectives do exist, even though they are still rather rare exceptions and often somehow remain "captives" of earlier stage mainstream scientific cultures, let us briefly look at Ashforth et al.'s (2008) review of scientific organizational corruption discourse. Calling for "theory development" that integrates "micro, macro, wide, long and deep views of organizational corruption" in view of a "considerably more holistic understanding", Ashforth et al. recognize that corruption (even though they still call it a dynamic "disease")"evolves in complex ways" (2008), and through "interacting subsystems" with their own inherent rationalities and subtle complexities and dynamics that our current level of theorizing has not yet captured. They therefore deplore that this kind of "deep view on corruption" is the one that is "least developed". Moreover, their assertion that "We management scholars/teachers might reasonably be asked, 'Are you part of the problem or part of the solution?' Arguably, we're both" indicates a degree of self-reflexivity which is rarely found in other publications. However, this short selective quote does not claim to be a systematic evaluation of the stage of reasoning complexity of Ashforth et al.'s article as a whole.

An even clearer meta-systematic perspective on corruption is presented by Alina Mungiu-Pippidi (2006) in her analysis of post-communist Romania (2006), claiming that "corruption can only be understood in conjunction with the stage of development of a particular state or society". In view of each society, she claims, "we must ask: are we dealing with modern corruption where corruption is the exception to the norm of universalism? Or are we dealing with particularism and a culture of privilege, where corruption itself is the norm? Or, as is frequently the case in the postcolonial world where the modern state was defectively implanted on a traditional society, are we dealing with a combination of the two? If so, to what extent is its main task to promote patronage and cater to specific interest groups?" Mungiu-Pippidi thereby clearly distinguishes different systems and their respective internal logics. To sum up, she concludes that "corruption will persist as long as human nature does not change. (…) An anticorruption agency might work well in democratic Australia, with its tradition of an independent judiciary, but the same kind of institution would fail to indict or arrest anybody who is 'somebody' in the former Soviet Union."

» DISCUSSION AND CONCLUSION

This paper was motivated by the desire to resolve a number of surprising, theoretically puzzling observations about present theorizing about corruption. More precisely, our primary theoretical motivation was to account for the plurality of strikingly different and sometimes even contradictory perspectives that important disciplines currently dealing with corruption within the social sciences take onto the phenomenon. We have asked how those differences and contradictions between and within disciplinary approaches can be explained and how they can eventually be bridged. While reviewing the literature, we also found substantially different understandings of corruption in different social and scientific cultures and contexts, as well as in different times, and asked how they can be explained and made sense of. Furthermore, we observed that mainstream western notions of corrupt and/or unethical behavior (which are usually taken for granted as "generally accepted" in large parts of the OB/BE/MS literature) are probably not representative in non-western context and are therefore not very helpful for dealing with corruption in many of those other contexts. In this regard, we asked to what extent those mainstream notions have to be contextualized – or at least to be made more explicit – in order to be able to make more generally valid claims about the phenomena in question. Finally, in view of the Model of Hierarchical Complexity as a new analytical tool presented here, we asked how it can help to solve the problems mentioned above, namely to integrate the perspectives, questions and findings from different disciplines and to thereby to provide a more complex and a more differentiated outlook on corruption/unethical behavior. In this regard, the MHC has in particular been compared to Kohlberg's classic model of the development of moral reasoning. So what are the main insights and contributions to be gained from this endeavor?

Theoretical contributions and implications

Our main contribution is to re-interpret phenomena of corruption, as well as of public, political and scientific ways of dealing with corruption in different times and different cultural contexts on the basis of an integrative model for analyzing physical and discursive behavior in a concise and non-arbitrary way, which can thus be considered as a universally applicable meta-systematic tool for detecting the structural patterns of both discourse and behavior. This general contribution can be broken down to methodological, empirical, and theoretical benefits and achievements, each of them showing why this kind of approach is an important improvement as compared to less complex perspectives.

Methodological relevance. Coming from the field of structuralist adult development theory and research, the Model of Hierarchical Complexity is based on content-free, mathematical analytical categories and definitions, which makes applicable in different disciplines and thus enables it to render views and findings from different theoretical and disciplinary backgrounds comparable. Providing tools for analyzing the complexity of reasoning and behavior it does not produce the usual culturally biased outcomes and can thus introduce more analytical rigor into the study of corruption in organizations.

Empirical relevance. Furthermore, our contribution is important for empirical reasons. Abundant research has shown that the development of cognition and other aspects of the mind does not stop after adolescence. With regard to moral development for example, it has been shown that "most adults are at the conventional level" of moral reasoning, and that "fewer than 20% of American adults reach the principled level (…), where actions should be more consistent with moral thought" (Treviño et al. 2006). This has strong, yet still largely underestimated consequences not only for social, political and economic life in general, and for organizational behavior in particular, but also for analyzing all of them. This observation expands on Treviño et al.'s (2006) claim that cognitive developmental dimensions have "clear implications for behavioral ethics in organizations". More precisely, recalling a statement by some of the leading adult development scholars cited earlier, organizations, as societies in general, "are comprised of individuals operating at multiple stages of development in various domains. Thus, political cultures and social systems display concurrent operations of several different stages. There are many overlapping systems and relationships among different people and entities" (Ross & Commons, 2008). At the same time, there are modal stages, i.e. stages at which most individuals operate within governments, societies, and organizations and which thereby characterize the stage at which the respective entities are likely to operate (Commons & Goodheart, 2007).

In other words, theories which don't take into account those empirical facts, fail to grasp an important dimension of behavioral reality and thus remain undercomplex. While many of the dominant research traditions tend to leave this dimension out of systematic analysis, adult development perspectives render them analyzable, and even place them into the center of analytical attention, thus offering an alternative, more complex way to frame intercultural corruption research.

Meta-systematic theory-integration. In view of theory integration, the theoretical and analytical power of the meta-systematic perspective presented here is able to account for corrupt behavior in different times, for different types of corrupt behavior, and for the differences in the attitudes of historical and present actors, societies and scientific discourses towards corruption at the same time (see Fein & Weibler, 2014). First, it integrates historical findings according to which corruption is a product of modernization during which public and private spheres came to be differentiated, and only as a result of which distinguishing between practices of and debates about corruption began to make sense. In fact, empirical descriptions of and findings about historically more distant societies prove to be compatible with lower levels of complexity development in general, as well as in present societies. Second, our meta-theory integrates the sociological finding that to what extent corrupt behavior comes to be critically reflected, depends to a large extent on variables like education and socio-cultural development. Moreover, in contexts where rather low levels of development in adults prevail, we tend to find higher levels of corruption (e.g., in developing countries). Third, our model integrates anthropological findings observing that on the behavioral micro level, practices of reciprocity, often considered as corrupt by higher stage reasoning if used in public, remain important not only on earlier levels of social development, but also in informal and private contexts in western societies. It thus makes clear that premature value judgments may prohibit an appropriate analysis of actual behavioral logics. Finally, our meta-theory is able to integrate questions and findings from many other disciplines in the fields studied in more detail here, namely Behavioral Ethics, Organizational Behavior and Management Studies.

As a result, the model proposed in this paper not only offers a better understanding of where ethical and unethical/corrupt behavior come from, addressing individual, organizational, and institutional influences on ethical behavior and, thereby, both micro, meso and macro levels of analysis. Because of its high degree of detachment and (self-) contextualization, it also offers a substantially new and more complex outlook on actor's understandings of and attitudes towards corruption. By making clear why perceptions differ not only between cultures, but also inside western societies, as well as inside scientific communities, it can bridge differences, gaps and contradictions in the literature and thereby reconcile previously disparate perspectives. As a structuralist framework, the MHC goes considerably beyond Kohlberg's model of moral development, for example by distinguishing more stages of complexity and by proposing even more objective criteria for their structuralist analysis. The approach proposed here can thus be situated in several academic literatures at the same time without being attached to any of them. It can therefore push research on behavioral ethics and on corruption in organizations into new paradigmatic insights and substantially advance the field in both theoretical and analytical respects.

Practical implications

Finally, besides the theoretical contributions mentioned above, our approach has also considerable practical implications. Meta-systematic, i.e. stage sensitive perspectives are able to more objectively and thus more efficiently adjust practical anti-corruption programs and activities to the respective nature of the problem. In other words, they define incentives, constraints, regulations and the like in view of the particular clientele they wish to serve. So instead of producing "one size fits all" solutions, they will come up with more differentiated strategies, depending on the level of complexity of reasoning and action of the concrete type of corruption in place, and of that of the actors trying to deal with it. For example, transparency regimes and ethics codes might work in systematic stage contexts while they will completely fail in connection with lower than formal stage actors, no matter if inside or outside western contexts. Unless a sufficient number of people in the context in question function on the basis of formal action logics, legal and bureaucratic solutions alone will not eliminate problems of corruption, but have to be combined with more traditional forms of authority. Similarly, appeals to ethics, fairness and social responsibility will not convince actors with less complex than systematic reasoning, because considering broader social consequences of their actions are not part of their reasoning structure. In this respect, our model clearly points out limitations of mainstream western strategies of corruption control "at home", but even more so in developing countries.

Limitations of the model and of the present paper

Of course, like every theory, ours equally has shortcomings and limitations. The same is true for the present paper. First, with regard to the meta-systematic perspective itself, its complexity of theorizing and perspective taking implies that it is ridden with prerequisites. One has to acquire a minimum familiarity with adult development theory and measurement techniques in order to be able to work with the model successfully. Second, when using the MHC for analyzing complex social phenomena such as organizations, one has to take into account the mutual influences of several levels of complexity of reasoning and action within the specific setting. While it may often appear that stages are not "pure", or difficult to identify empirically, task definitions have to be carefully defined and constructed for each new study. Furthermore, larger empirical studies based on the MHC are rather intricate and time-consuming, since to achieve high inter-rater reliability, raters have to be intensively trained in working with the model. However, this concern applies to any new, notably complex measurement method and must therefore not be stacked against it. Finally, depending on the context in which the MHC, or more generally, adult development theory, shall be used, it has to be taken into account that in some parts of the social sciences, it is not easily accepted to claim developmental differences between social actors or cultures – even though this critique often seems to come from a spontaneous (and rather superficial) reflex rooted in specific scientific cultures, rather than from a thorough examination of the model itself. However, similar criticisms can probably be countered successfully based on results of the model's application in a certain field. At the same time, while the MHC offers a way of better understanding differences in development, as well as positions criticizing them, the model itself is, of course, a theoretical lens which can also be contextualized, and might eventually be evaluated on the basis of either more complex perspectives or otherwise well founded arguments at some point.

If space permitted, it would have been desirable to give more detail on how to use the MHC in concrete research settings, to supply more examples for corrupt phenomena and situations as seen, scored and explained by the model in order to make the theoretical and analytical gains offered yet more convincing. Even though we did spell out some important practical implications of our approach and provided a number of empirical examples, a more comprehensive discussion of applying the MHC on different aspects of corruption research was beyond the limits of this paper. Despite these caveats, the present paper demonstrated important theoretical and, at the same time, meta-theoretical contributions to be gained by a more systematic use of adult development perspectives on corruption. ∎

REFERENCES

Aguilera, R.V., & Vadera, A.K. (2008). The dark side of authority: Antecedents, mechanisms, and outcomes of organizational corruption. Journal of Business Ethics, 77(4), 431–449. doi: 10.1007/s10551-007-9358-8

Ashforth, B. E., Goia, D. A., Robinson, S. L., Trevino, L. K. (2008). Introduction to special topic forum: Re-viewing organizational corruption. Academy of Management Review, 33(3), 670–684. doi: 10.5465/AMR.2008.3246

Beugré, C. D. (2010). Resistance to socialization into organizational corruption: A model of deontic justice. Journal of Business Psychology, 25(3), 533–541. doi: 10.1007/s10869-010-9176-3

Chang, J. J., & Lai, C. C. (2002). Is the efficiency wage efficient? The social norm and organizational corruption. Scandinavian Journal of Economics 104(1), 27-47. doi: 10.1111/1467-9442.00270

Commons, M. L. (2008). Introduction to the model of hierarchical complexity and it relationship to postformal action. *World Futures*, *64*(5-7), 305-320. doi: 10.1080/02604020802301105

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. (2009). Further issues in stage metrics. *Integral Review: A Transdisciplinary and Transcultural Journal for New Thought, Research, and Praxis, Vol. 5* (2), 315-323.

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(2), 247-267. doi:10.1080/03057240600681785

Commons, M.L., Gane-McCalla, R., Barker, C. & Li, E.Y. (2014). The Model of Hierarchical Complexity as a Measurement System. *Behavioral Development Bulletin, 19*(3).

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*(3-4), 91-111. doi: 10.1007/s10804-007-9016-5

Commons, M. L., & Richards, F. A. (1984a). A general model of stage theory. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), *Beyond formal operations: Vol. 1. Late adolescent and adult cognitive development* (pp. 120-140). NewYork: Praeger.

Commons, M. L., & Richards, F. A. (1984b). Applying the general stage model. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), *Beyond formal operations: Vol. 1. Late adolescent and adult cognitive development* (pp. 141-157). NewYork: Praeger.

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, Vol. 10*, 323–340.

Commons, M. L. & Rodriguez, J. A. (1993). The development of hierarchically complex equivalence classes. *Psychological Record, 43*(4), 667–697.

Commons, M. L., Sinnott, J. D., Richards, F. A., & Armon, C. (Eds.) (1989). *Adult development: Comparisons and applications of developmental models: Vol. 1*. Westport, Connecticut: Praeger.

Commons, M. L., Armon, C., Kohlberg, L., Richards, F. A., Grotzer, T.A., & Sinnott, J.D. (Eds.) (1990). *Adult development: Models and methods in the study of adolescent and adult thought: Vol. 2*. New York: Praeger.

Commons, M. L., & Sonnert, G. (1994). Society and the highest stages of moral development. *Politics and the Individual, 4*(1), 31-55.

Commons, M. L., Trudeau, E., Stein, S., Richards, F., & Krause, S. R. (1998). Hierarchical complexity of tasks shows the existence of developmental stages. *Developmental Review, 18*(3), 237-278. doi:10.1006/drev.1998.0467

Coombs, C. H., Dawes, R. M., & Tversky, A. (1970). *Mathematical psychology: An elementary introduction*. Englewood Cliffs, New Jersey: Prentice-Hall.

Dawson-Tunik, T. L. (2006). The meaning and measurement of conceptual development in adulthood. In C. Hoare, (Ed.), *Handbook of adult development and learning* (pp. 433-454). New York: Oxford University Press.

Fein, E. (2012). Kognition und politische Kultur. Sozialpsychologische perspektiven in der diskursforschung am beispiel von patronage, klientelismus und korruption in Russland. Cognition and political culture. Adult development perspectives on patronage, clientelism and corruption in Russia] In Oberthuer, J., Rosa, H. (Eds.), *Normen- und Wertbegriffe in Europa im Wandel* (in print).

Fein, E., & Weibler, J. (2014). Cognitive basis for corruption and attitudes towards corruption viewed from a structuralist adult developmental meta-perspective. *Behavioral Development Bulletin, 19*(3).

Fleck, C. & Kuzmics, H (Eds.) (1985). *Korruption: Zur soziologie nicht immer abweichenden Verhaltens*. [Corruption: Sociology of not always deviant behavior] Königstein/Ts.: Athenäum.

Kegan, R. (1982). *The evolving self: Problem and process in human development*. Cambridge, Massachusetts: Harvard University Press.

Lange, D. (2009). A multidimensional conceptualization of organizational corruption control. *Academy of Management Review, 33*(3), 710—729. doi: 10.5465/AMR.2008.32465742

Ledeneva, A.V. (1998). *Russia's economy of favours*: *Blat, networking, and informal exchange*. Cambridge: Cambridge University Press.

Levinson, D.J. (1986). A conception of adult development. *American Psychologist, 41* (1), 3-13. doi: 10.1037/0003-066X.41.1.3

Lindsay, P. H., & Norman, D. A. (1977). *Human information processing: An introduction to psychology* (2. ed.). New York: Academic Press.

Luo, Y. (2004). An organizational perspective of corruption. *Management and Organization Review 1*(1), 119—154. doi: 10.1111/j.1740-8784.2004.00006.x

Maravic, P. V. (2007). Decentralized corruption in Germany. *Public Administration Quarterly*, *30*(4), 440-478.

Maravic, P. V. (2007b). Public management reform and corruption: Conceptualizing the unintended consequences. *Administration and Public Management Review, 8*, 126-148.

Marchand, H. (2002). Some reflections on post-formal thought. *The Genetic Epistemologist*, *29* (3).

Merl, S. (2012). Gibt es eine spezifisch russische wirtschaftskultur? Reflexionen über die administrative kommandowirtschaft und ihre auswirkungen bis heute. [Is there a specific Russian economic culture? Reflections on the administrative commando economy and its implications until today] *Geschichte und Gesellschaft, 38*, 81-109.

Merl, S. (2010). Kann der korruptionsbegriff auf Russland und die sowjetunion angewandt werden? [Can the term corruption be applied to Russia and the Soviet Union?] In N. Grüne, & S. Slanicka (Hg.), *Korruptio*: *Historische annäherungen an eine grundfigur politischer kommunikation* (pp. 247-279). Göttingen: Vandenhoeck & Ruprecht.

Merl, S. (2008). Die korruption in Russland heute - ein Vermächtnis Stalins? In C.Söller, & Wünsch, T. (Eds.), *Korruption in ost und west: Eine debatte* (pp. 33-78), [Corruption in east and west: A debate], Passau: Stutz.

Mungiu-Pippidi, A. (2006). Corruption: Diagnosis and treatment. *Journal of Democracy, Vol. 17* (3), 86-99. doi: 10.1353/jod.2006.0050

Paul, A. T., & Schwalb, B. (2011). Wie organisiert ist das organisierte verbrechen? Warum es die mafia nicht geben dürfte und warum es sie trotzdem gibt. [How organized is organized crime? Why mafia should not exist, but does still exists] *Leviathan, 39*, 125—140. doi: 10.1007/s11578-011-0112-3

Pelletier, K. L., & Bligh, M. C. (2006). Rebounding from corruption: Perceptions of ethics program effectiveness in a public sector organization. *Journal of Business Ethics, 67*(4), 359—374. doi: 10.1007/s10551-006-9027-3

Rabl, T. (2011). The impact of situational influences on corruption in organizations. *Journal of Business Ethics*, *100*(1), 85—101. doi: 10.1007/s10551-011-0768-2

Ross, S. (2004). *Roots and manifestations of patronage and clientelism*. Unpublished manuscript.

Ross, S. (2007). A developmental behavioural analysis of dual motives' role in political economies of corruption, paper presented in *Dual Motive Theory as applied to economic and political development*, New York.

Ross, S., & Commons, M. L. (2008). Applying hierarchical complexity to political development. *World Futures, 64*(5-7)*,* 480—497. doi: 10.1080/02604020802303721

Schattenberg, S. (2008). *Die korrupte provinz. Russische beamte im 19. jahrhundert*. [The Corrupt Province. Civil Servants in Russia in the 19th Century]. Frankfurt/New York: Campus Verlag.

Shadnam, M., & Lawrence, T. B. (2011). Understanding widespread misconduct in organizations: An institutional theory of moral collapse. *Business Ethics Quarterly, 21*(3), 379-407.

Treviño, L. K. (1986). Ethical decision making in organizations: A person situation interactionist model. *Academy of Management Review, 11(3),* 601-617. doi: 10.5465/AMR.1986.4306235

Treviño, L. K., Weaver, G. R., & Reynolds, S. R. (2006). Behavioral ethics in organizations: a review. *Journal of Management* 2006, *32*(6), 951-990. doi: 10.1177/0149206306294258

Venard, B., & Hanafi, M. (2008). Organizational isomorphism and corruption in financial institutions: Empirical research in emerging countries. *Journal of Business Ethics, 81*(2), 481—498. doi: 10.1007/s10551-007-9519-9

Voliotis, S. (2011). Abuse of ministerial authority, systemic perjury, and obstruction of justice: Corruption in the shadows of organizational practice. *Journal of Business Ethics, 102*(4), 537-562. doi: 10.1007/s10551-011-0826-9

Voslensky, M. S. (1987). *Nomenklatura*: *The* Soviet *ruling class*. New York: Doubleday.

BEHAVIORAL DEVELOPMENT BULLETIN

Measuring care-based moral development: The ethic of care interview

Eva E. A. Skoe

University of Oslo

ABSTRACT

This paper presents a recently developed instrument of care-based moral development: The Ethic of Care Interview (ECI) (Skoe, 1998, 2008). Based on Carol Gilligan's (1982) theory, the ECI measures five levels of care-based moral thought. These range from an initial position of self-concern, through questioning of self-concern as a sole criterion; to a position of primarily other-concern, questioning of other-concern as a sole criterion; and finally balanced self and other concern. The stages involve a progressively more complex understanding of human interdependence and an increasing differentiation of self and other. The semi-structured ECI interview consists of a real-life moral conflict generated by the participant and three standardized dilemmas. Administration and scoring as well as reliability and validity are described. A series of studies has shown that balanced consideration of the needs of self as well as others appears to develop gradually across childhood into young adulthood. Research findings point to the importance of care-oriented morality for human growth, especially identity and personality development. Further research with the ECI is suggested.

KEYWORDS: care-based moral development, ethic of care interview

Essentially, the ethic of care reflects a cumulative understanding of human relationships based on the recognition that self and other are interconnected. Just like violence or harm leads to destruction, caring lead to benefitting both self and the others (Gilligan,1982). Since Gilligan's original critique of gender bias in Kohlberg's (e.g., 1984) work on justice-based moral reasoning, there has been considerable controversy and debate regarding possible sex differences in moral development. Gilligan observed that when women were asked to talk about their personal real-life moral dilemmas they often described issues of care and responsibility in relationships that were not well described in Kohlberg's justice-oriented model. The failure of women to fit those models of human growth, Gilligan (1982) argued, may point to a problem not in women's development, but "in the representation, a limitation in the conception of human condition, an omission of certain truths about life" (p. 2). In her view: the ethic of justice best represents the moral reasoning and values of men, whereas the ethic of care represents the moral reasoning and values of women. The empirical evidence on sex differences in moral reasoning is quite complex and controversial. Overall, it seems both men and women have both justice and care orientations available and use them differentially depending on various background and contextual factors (Jaffe & Hyde, 2000; Pratt, Skoe & Arnold, 2004). In their meta-analysis of this literature Jaffe and Hyde (2000) found some modest differences in support of the hypothesis that women are higher in an orientation toward care-related moral issues, whereas men are higher in an orientation to justice. In particular: women are more likely to discuss real-life moral conflicts involving close personal relationships than men. Women also have been observed to view a variety of dilemmas as more important and more difficult to resolve than do men. Thus, there is evidence that women and men differ in terms of how they experience and evaluate moral situations (Skoe, Cumberland, Eisenberg, Hansen, & Perry, 2002; Skoe, Eisenberg, & Cumberland, 2002; Wark & Krebs, 1997).

Gilligan has been credited with extending the moral domain to include a moral orientation of care (i.e., concern with responsibility, harmony and prevention of hurt in interpersonal relationships) as well a moral orientation of justice (i.e., concern with equality, fairness and individual rights). Today it is generally acknowledged that thinking about care, context and relationship issues is an important component of morality. This has drawn attention to Gilligan's work (e.g., Walker, 2006).

The extensive focus on sex differences has, however, tended to obscure another important implication of Gilligan's (1982) theory: care reasoning, like justice reasoning, follows specific developmental pathways and varies individually. Based on the one-year follow up of 21 women, ranging in age from 15 to 33 (the Abortion Decision Study), she discussed a developmental continuum in their care orientation. Skoe (1998, 2013) has explored Gilligan's suggestion of developmental trends in the growth of care-oriented moral thought in both men and women, through the construction and validation of the Ethic of Care Interview (ECI). The ECI appears to be the first and, to my knowledge, only attempt at operationalizing Gilligan's theory regarding the developmental aspects of the care ethic.

The purpose of the ECI is to locate individuals in one of the Ethic of Care levels based on their responses to four moral dilemmas. Following Gilligan (1982), each level represents a different mode of resolving conflicts in human relationships and a different apprehension of the central concept that self and other are interdependent. The five levels involve a progressively more complex understanding of human relationships and an increasing differentiation of self and other. These levels will be discussed below. The ECI provides a tool with which to answer questions about the relevance and usefulness of care-based approach to moral thought (Skoe & Marcia, 1991).

» THE ETHIC OF CARE INTERVIEW
ECI levels and sample responses

The ECI consists of four dilemmas administered in a semi-structured interview format. In addition to a real-life conflict generated by the participant, three standard interpersonal dilemmas are presented that involve conflicts about (a) unplanned pregnancy, (b) marital fidelity, and (c) care for a parent (see Appendix). In line with the theories of Haan (1975) and Gilligan (1982), these dilemmas are used because they represent frequently occurring situations of interpersonal concerns where helping others could be at the price of hurting oneself.

Based upon an initial pilot study with women, and subsequent studies with men and women (e.g., Skoe, 1986; Skoe & Marcia, 1991; Skoe & Diessner, 1994), the ECI manual (Skoe, 1993) was constructed containing descriptions congruent with Gilligan' theory (1982) and sample responses for five ethic of care levels. The care levels involve moving from an initial position of self-concern, through questioning of self-concern as a sole criterion; to a position of primarily

other-concern, questioning of other-concern as a sole criterion; and finally balanced other and self concern. The three primary care levels, and the two transitional levels, are the following:

Level 1: Survival (caring for self); the lowest level in the ECI sequence, individuals think about relational issues in self-protective, pragmatic way, and neglect the needs or feelings of others. The aims are basically to ensure one's own happiness and to avoid pain. There is no consideration of abstract ethical principles or values. The following are brief descriptions of the ECI levels and sample responses to the Betty/Erik dilemma (see Appendix) about an unhappy marriage to a recalcitrant spouse and the possibility of an emotionally satisfying extramarital relationship:

I don't think he should keep having an affair on his wife because that is going to end up not going in a positive direction. (Why shouldn't he have an affair?) Either his mistress is going to want him to leave his wife or his wife is going to find out. He is just going to get himself into more problems ... I think life is too short to stay in an unhappy situation. There are too many other opportunities to be happy. (Why is it important to be happy?) We are only here for about 80 years or so, we may as well make the best of it.

Level 1.5 concerns the transition from self-care (survival) to a sense of responsibility. Concepts of selfishness and responsibility first appear at this level. Caring for the self to ensure survival is criticized as selfish. In relationships, although one may be aware of the needs of others, one gives more importance to one's self-interest.

There's actually three angles you can take it from. The first one would be Erik's happiness. If he's unsatisfied, he should do it. But from a legal/financial standpoint he shouldn't do it because he'd get screwed in the end, just like Derek would, and he'd get into that dilemma where he would lose his kids and Betty would divorce him, he'd lose a lot of money, Carol might leave ... Religiously, I don't know, slash morally, I guess, he shouldn't do it, just because he's married ... It would be very selfish. If you're unhappy try to fix the situation. If not, get a divorce quickly.

Level 2: Caring for others; individuals reason about issues in terms of responsibility and care for others to the exclusion of the needs of self. Being *good* is equated with self-sacrificial concern for other people, and what is *right* is externally defined, often by the parents, church, or society. There is a strong need for security. Being accepted or liked by other people is so important that others may be helped and protected, even at the expense of self-assertion.

I don't believe in divorces or extramarital flings. She could try other ways to make her husband realize that she wants a bit more out of the marriage, possibly volunteer work or take a part-time job. The kids are old enough to be left alone some of the time... She has been married a long time. She should try a bit harder to get through to her husband. She has children, divorce is hard on children. I believe in marriage and staying together. Marriage is a commitment, you should stay married.

Level 2.5 concerns transition to a reflective care perspective, marked by a shift in concern from goodness to truth and personal honesty in relationships. Compared to the more "black-and-white" worldview of the previous level, complexities and nuances are expressed. The goodness of protecting other people at one's own expense is questioned.

Communication doesn't seem to be too good between her and her husband. Her happiness is important because it affects the way you raise your children. If you're not happy in a situation I think you should resolve it. Maybe she should tell her husband that she likes someone else now, or, I guess, divorce or something like that. Whichever way she feels she is more confident about

herself... I think it has a big influence on the kids. Divorce would as well. But if you weigh out the two, an unhappy marriage could be worse for the kids. If he is not going to listen, obviously she does not have a good relationship. You can't have a family if you can't communicate to each other. I think it is best that she get out of it then, put herself into a family where she is more settled and relaxed and the communication is better.

Level 3, individuals fully realize the ethic of care (caring for both self and others). The needs and welfare of both others and self are encompassed in a more balanced approach to thinking about relationships. The tension between selfishness and responsibility is resolved through a new understanding of human interconnectedness. Out of this realization, the insight arises that by caring for others, you care for yourself, and vice versa; compassion enriches both the giver and the receiver. Concern is expressed for everyone impacted in the situation, and attempts are made to minimize hurt to all parties.

I think that he should seek counseling personally and possibly try and get his wife in some type of counseling as well. I think in this relationship there is more at stake, as they have two children which is a big concern. I don't have children, but I assume that I will have a very strong bond with my children and I would not want to do anything to hurt that. So my advice would be to seek professional help from people who are experienced in dealing with situations like these on a daily basis… If that didn't work, I would seriously consider divorce, if the situation was bad enough. I couldn't live in a miserable situation like that for an extended length of time because I feel that it would just deteriorate to arguing all the time or just a cold indifference, and I don't think either situation is good or beneficial for either the wife or the husband or the children.

Administration and scoring

The ECI is a semi-structured, individually administered interview which takes about 30 minutes (from 15 to 45 minutes) to complete. To avoid biasing real-life choice by providing an example beforehand, the real-life dilemma is elicited first in various ways: "Have you ever been in a situation where you were not sure what was the right thing to do?" "Have you ever had a moral conflict?" "Could you describe a moral conflict?" Adapted from the work of Gilligan (1982), these questions eliciting a dilemma are then followed by a set of six probe questions: "Could you describe the situation?" "What were the conflicts for you in that situation?" "In thinking about what to do, what did you consider?" "What did you do?" "Did you think it was the right thing to do?" and "How do you know?" The standardized dilemmas are read aloud to the participants while they read along. Probes such as "What do you think Betty/Erik should do?" and "What would you do if you were in the same situation? Why?" are used to examine dilemma reasoning. The interviews are audio taped for later transcription and scoring. Transcription is not always necessary; the interviews also can be scored from listening to the tapes. Scoring an ECI tape takes about the same time as the actual interview; a transcript takes less time to score.

In determining the level of a person's response, it is important to note whose needs and concerns the person considers in the dilemma situations, and the reasons *why* s/he would or would not do or say something. *What* the person would do is of lesser importance. For example, in the dilemma cited above, a person would be assessed at ECI level 2 (caring for others) either if thinking that Betty/Erik should stay married or if thinking that Betty/Erik should divorce if the reason given primarily is that "it is better for the children". In each dilemma, the person should be given ample opportunity to express her or his views and values on each dilemma without the help of suggestions from the interviewer. Conducting a good interview requires both practice and sensitivity (Skoe, 1993).

The Ethic of Care Interview can be scored according to total score across the four dilemmas, yielding a potential range of 4.00–12.00 for any single participant, or according to level, yielding five discrete levels. Based on the interview, the participant is given a level score for each dilemma. Quarter scores (e.g., 1.75, 2.25) can be assigned on any given dilemma if the response appears to fall between two levels, but should be used sparingly. If the person does not generate a real-life dilemma, the mean score for the other three dilemmas may be used in place of a real-life score.

Total scores are calculated by summing the ratings on the four dilemmas. Overall level scores on the Ethic of Care Interview are determined by dividing the total scores by four and then rounding to the nearest .5 level (e.g., 1.15 = Level 1; 2.45 = Level 2.5; 2.80 = Level 3). If a person's overall level score falls exactly between two levels (e.g., 2.25, 1.75), a second rater independently scores the person at one of the two adjacent levels.

Reliability

With regard to inter-rater reliability, a difference between two raters no greater than quarter of a level score (e.g., 2.50 and 2.75, is considered agreement; 2.50 and 3.00 is considered disagreement). Correlations between trained raters generally have ranged from .85 to .95 (Cohen's Kappa .86–1.00). It appears that training is not always necessary; acceptable inter-rater reliability (.78–.91, Kappa .63–.91) also has been obtained between an untrained rater and trained raters (Skoe & Marcia, 1991). Some self-training or practice in interviewing and scoring according to the ECI manual (Skoe, 1993) is, however, strongly recommended before undertaking research with the ECI.

Inter-correlations among the four ECI dilemmas commonly have ranged from about .70 to .90, and correlations of each dilemma with the total score have ranged from .73 to 97. Cronbach's alphas from .86 to .97 also have been calculated. Hence, it appears that the ECI can be scored with a fair degree of inter-rater reliability and internal consistency.

Concurrent validity

The ECI and justice-based moral reasoning. Since the ECI is a measure of care-based moral reasoning, proposed as an alternative to justice-oriented moral reasoning (Gilligan, 1982) a person's scores on the ECI should be positively correlated with justice reasoning tests. The models of both Kohlberg and Gilligan have a basis in cognitive developmental (i.e., Piagetian) theory. Re-

search supports this expectation. In the initial study on women (Skoe, 1986; Skoe & Marcia, 1991), there was a positive correlation between the ECI and the Sociomoral Reflection Measure (SRM), a written version of Kohlberg's Moral Judgment Interview (MJI; Colby & Kohlberg, 1987), developed by Gibbs and Widaman (1982), $r(86) = .37$, $p < .001$. In the subsequent Skoe and Diessner (1994) study, partial correlations controlling for age showed significant positive relationships between the ECI and the MJI both for women, $r(73) = .25$, $p < .02$, and for men, $r(55) = .34$, $p < .01$. In a later study by Skoe and Lippe (2002), the correlation between the ECI and Rest's (1979) multiple-choice justice instrument, the Defining Issues Test (DIT), was also positive, $r(141) = .21$, $p < .05$. Similarly, in a more recent investigation, Juujärvi, Myyry, and Pesso (2010) observed a correlation of $r(116) = .31$, $p < .01$, between the ECI and post-conventional scores as measured with the DIT.

The positive relationship between the ECI and justice-based moral reasoning as measured by the MJI also has been replicated by other studies. In a Finnish longitudinal study of nursing, social work and law enforcement students (assessed at the beginning of their studies, Time 1, and after two years of studying, Time 2), ECI and MJI scores were highly related at both times on real-life dilemmas, $r_1(57) = .78$ and $r_2(59) = .72$, as well as on hypothetical dilemmas, $r_1(57) = .65$ and $r_2(59) = .53$, all $ps < .001$ (Juujärvi, 2006). Similarly, in a sample of middle-aged and older Canadian adults, the correlation between the ECI (calculated on the basis of two real-life dilemmas) and MJI scores was $r(33) = .55$, $p < .01$ (Skoe, Pratt, Matthews, & Curror, 1996, Study 2). These findings indicate that the care and justice systems, although focusing on somewhat distinctive conceptions of morality and moral duties, as argued by Gilligan (e.g., 1982), may share underlying general developmental processes, perhaps in terms of role-taking (Skoe et al., 1996) or ego development (Skoe & Lippe, 2002).

Construct validity

The ECI and ego identity. Theoretically, there is a close connection between morality and ego identity. Both are assumed to be related to cognitive development, involving similar processes such as conflict or disequilibration, exploration and commitment (Marcia, Waterman, Archer, & Orlofsky, 1993). Whereas Kohlberg (1973) believed that certain features of ego development are a necessary but not sufficient condition for the development of moral structures, Marcia (1980) speculated that identity and moral reasoning are related reciprocally. Kohlberg and Gilligan (1972) wrote that "Erikson's picture of an adolescent stage of identity crisis and its resolutions… is a picture dependent upon formal logical thought and of questioning conventional morality" (p. 1078). Accordingly, there should be a positive relationship between the ECI and ego identity development. In the initial study of 86 Canadian university women, 17–26 years of age (Skoe, 1986; Skoe & Marcia, 1991), the ECI was positively related to age, $r(84) = .44$, ego identity development (Marcia et al., 1993), $r(84) = .86$, and to the SRM, $r(84) = .37$, all $ps < .001$.

Because the above study was restricted to women, it could not address possible sex differences or whether care-based moral thought is more applicable to women than to men, as Gilligan (1982) argued. Therefore, a subsequent study by Skoe and Diessner (1994) of 58 men and 76 women, all university students in USA, 16–30 years of age, was conducted, examining the relations among ego identity, care-based and justice-based moral reasoning with use of Kohlberg's interview (MJI). This extended investigation showed that the ECI was positively related to age for both men, $r(56) = .30$, $p < .02$, and women, $r(74) = .52$, $p < .001$, and to justice-based morality (see above) as well as strongly related to ego identity. Controlling for age, the partial correlation between the ECI and identity was $r(73) = .78$, for women, and $r(55) = .59$, for men, both $ps < .0001$.

The very high correlations between ECI and identity for women in the forgoing two studies ($rs = .86$ and $.78$) are likely due to both involving a similar underlying process of the thoughtful exploration of alternatives, even though the interview content for identity statuses (life domains) and ethic of care (relational dilemmas) are different. The former refers to one's own life, the latter to solutions to problems of others. It is the underlying *processes* that seem to be similar. That the correlations for women are higher than those for men may be due to the greater relevance of the *content* of the ECI to women's identity. Clearly, because the two are labeled differently, a discriminant validity study would be in order. This might take the form of an incremental validity investigation in which both variables could be assessed for their relative contribution to common broader variables such as ego development or Common's (2008) Model of Hierarchical Complexity, or to a behavioral variable such as moral action in a structured situation.

There were no significant sex differences on the identity, justice, or care measures themselves. Although these results suggested that sex differences may not be as pronounced as proposed by Gilligan (e.g., 1982), further analyses indicated the care ethic nevertheless may operate differentially in men and women in important ways. For example, the partial correlation between the ECI and identity was significantly higher than the correlation between the MJI and identity for women only, $t(73 = 6.05)$, $p < .001$ (for further details see Skoe & Diessner, 1994). Hence, the ethic of care may be a more central component of ego identity for women than for men. Replicating the results of Skoe and Marcia (1991), for women care-based moral development was more highly related to identity than was justice-based moral reasoning. These results are consistent with Gilligan's (1982) argument that women's conceptions of self and morality are intricately linked and that the care ethic has special relevance for women's personality development. The care ethic may influence women's everyday life experiences and thought more than men's, perhaps due to culture, socialization and activity preferences (Skoe, 1998).

The ECI and ego development. Further construct validity was obtained by relating the ECI to ego development. Conceptually: morality has been considered an aspect of ego development (e.g., Blasi, 1998; Gilligan, 1982; Kohlberg, 1984; Loevinger, 1979). In Loevinger's view: ego is that aspect of personality that establishes a basic unity by constructing the meanings one gives to oneself, to other people and to the social world. Her sentence completion test measures sequential stages in the growth of this broad construction of meaning. Empirically: a positive link between ego development as measured with Loevinger's model and justice

reasoning is reasonably well established (e.g., Snarey, 1998). A fairly recent Norwegian study by Skoe and Lippe (2002) examined the relations among ego development and justice and care ethic reasoning levels in 72 men and 72 women, 15 to 48 years old, considering age, education, sex, and verbal intelligence.

As predicted on the basis of theory and previous research (e.g., Kohlberg, 1984; Loevinger, 1979; Skoe, 1998; Skoe & Diessner, 1994), both care as measured by the ECI and justice as measured with the DIT (Rest, 1979) were positively related to ego development as measured with the Washington University Sentence Completion Test (SCT; Loevinger & Wessler, 1970; Hy & Loevinger, 1996). Furthermore, the correlation between ECI total scores and ego development, $r(142) = .58$, $p < .0001$, was significantly higher than the one between DIT P scores and ego development, $r(139) = .20$, $p < .02$, $t(138) = 4.38$, $p < .0001$. When age, education, and verbal intelligence were controlled the relation between ego development and ECI remained significant, $r(136) = .51$, $p < .001$, but the relations between ego development and the DIT, $r(136) = .13$, ns, and between the ECI and the DIT, $r(136) = .13$, ns, were not.

The ECI's low positive relationship to verbal intelligence, $r(142) = .34$, $p < .001$, in this study, is reasonable considering it is a verbal test assessing complex reasoning about care for self and others. The results also demonstrate, however, that the ECI is conceptually distinct from verbal intelligence. With verbal intelligence partialed out, the variance shared between the ECI and ego development remained substantial: $r(141) = .53$, $p < .001$. By contrast, the DIT did not relate significantly to ego development or to care reasoning, $rs(138) = .11$, ns, and $.11$ ns, when the effect of verbal intelligence was controlled. Hence, the results suggested that the weak positive relationships between the DIT and ego development, as well as between DIT and ECI are reduced to their common overlap with verbal ability, likely "the most salient marker of general intelligence" (Sanders, Lubinski, & Benbow, 1995, p. 502). This is not the case for the relation between the ECI and ego development (Skoe & Lippe, 2002). A principal component factor analysis with a promax (oblique) rotation was computed for the major variables in this study. Three factors were extracted with eigen values greater than one that accounted for 41% of the total factor variance. All variables loaded on one of three factors as follows: Factor 1 – SCT .87, and ECI .86; Factor 2 – age .88, and education .77; Factor 3 - DIT .90, and verbal intelligence .63.

Ego and care development appear to have more in common with each other than they have with justice development. This may be the case especially at the higher stages where both ego and care developments involve greater ability to integrate respect for personal autonomy with responsibility, compassion, and intimacy in relationships. Therefore, ego development and care development may be seen as mutually enhancing.

The ECI and empathy. Another evaluation of construct validity was the ECI's relationship to empathy. Conceptually: there is a close connection between care and empathy. Empathy is generally viewed as a multidimensional construct, involving three separate dispositions: perspective taking, empathic concern (or sympathy), and personal distress. To see if these dispositions might be differentially related to care-based moral development the ECI and Davis's (1996) multidimensional empathy measure were administered to 58 Canadian university students (30 were women), ranging in age from 20 to 42 years (Skoe, 2010). Partial correlations, controlling for age, parents education, and sex, showed that empathic perspective taking was positively (.37), and personal distress negatively (-.36) related to ECI scores, both p's < .01. Analyses also indicated a curvilinear relationship between the ECI and sympathy for women, but not for men. Women at ECI Level 2 (self-sacrificing care for others) scored significantly higher on sympathy than did all others. This was predicted because sympathy involves high emotional reactivity and more selfless concern for others (Davis, 1983, 1996), paralleling the "other-oriented" position of ECI Level 2.

In sum, the results of this study support the view that empathy plays a constructive role in care-oriented moral development (e.g., Hoffman, 2000). Participants who demonstrated more integrated care reasoning also showed greater tendencies to see the world from others' points of view as well as lower levels of anxiety and uneasiness in reaction to others' distress. However, the causal ordering of these relationships cannot be determined from these correlational data, and further research is required. The relations between care reasoning levels and empathy-related responding likely are bi-directional. Dispositional perspective taking could be underlying the capacity to consider mutually the needs of others and self as assessed on the ECI, but it also is plausible that higher levels of care reasoning provide the potential for people to understand or consider others' situations and viewpoints. In any case, the findings provide some convergent and discriminant validity for the ECI. As always, because the sample is homogenous and modest in size, replication is required. In a Finnish sample of with students in various fields (Juujärvi et al., 2010), the ECI was positively associated with both empathic perspective taking and meta-ethical thinking, rs (127) = .29 and .31, ps < .01. Furthermore, the ECI and sympathy were positively related for men, $r(35) = .40$, $p < .01$, but not for women, whereas ECI and personal distress was negatively related (marginally) for women, $r(95) = -.20$, $p = .051$, but not for men. Further cross cultural studies are required; both care reasoning and empathy may be differentially valued or emphasized for women and men across time and nations.

The ECI and social interaction. With regard to intimate social interaction, recently Skoe, Pratt and Øvregård (2011) examined the links among care-based moral development, commitment, and trust measured with the Trust in Close Relationships Questionnaire (TIR; Rempel, Holmes, & Zanna, 1985) in 90 Canadian young adults, approximately 26 years of age. As predicted: participants who were in a committed romantic relationship scored significantly higher both on care reasoning and interpersonal trust, compared to those who were not in a committed relationship. Care reasoning and trust were significantly correlated (.30, $p < .01$). A mediation analysis (Baron & Kenny, 1986), however, suggested that this relationship was partly mediated by being in a committed relationship (Sobel test = 2.28, $p < .03$). These findings suggest that people higher in the ECI may be more able to establish and maintain a committed romantic relationship, which in turn might lead to higher levels of interpersonal trust.

Higher ECI levels also are linked to greater volunteer participation such as helping sick and elderly people as well as charity donations (Pratt et al., 2004; Skoe, 1998). By contrast: 51 forensic

psychiatric patients in mid adulthood, all who had committed violent acts, such as homicide and rape (Adshead, Brown, Skoe, Glover, & Nickerson, 2008), with the exception of two persons, all scored at the lowest ECI levels (survival, caring for self). These data suggest that the ECI is associated positively with prosocial behavior and negatively with antisocial or criminal behavior.

The ECI, age and education. Both cross-sectional and longitudinal data show that care reasoning levels as tested with the ECI are positively related to age in adolescents and young adults, but relatively stable in mid- to late adulthood. This is congruent with findings on justice-based moral thinking (e.g., Juujärvi, 2006; Pratt et al., 2004; Skoe et al., 1996), and thereby provide some construct validity for the ECI as a developmental measure.

Whereas it has been fairly well documented that justice reasoning is positively related to education (e.g., Skoe et al., 1996), research suggest that care reasoning depends inconsistently on this variable. Some studies show non-significant correlations between ECI scores and level of education, for example, women, $r(28) = .17$, *ns*, and $r(19) = .21$, *ns*, men, $r(28) = .16$, *ns*, and $r(14) = -.01$, *ns* (Skoe et al., 1996, Study 1 and 2 respectively); women, $r(63) = -.04$, *ns*, men $r(27) = .13$, *ns* (Skoe et al., 2011). Others show small to moderate positive correlations, for example, women, $r(72) = .32$, $p < .01$, men, $r(72) = .24$, $p < .05$, or combined, $r(144) = .27$, $p < .005$ (Skoe & Lippe, 2002). This inconsistent pattern of results, which also is the case for ego development (e.g., Hauser, 1976), should be examined more closely. The effects of education might depend on length, type and quality, for instance, and might vary across age and nations. For example, Juujarvi (2006) observed that over a 2-year period, Finnish social work students progressed in care reasoning, but law enforcement students did not. Moreover, perhaps not surprisingly, people studying social work scored higher on the ECI than people studying business (Juujärvi et al., 2010).

The ECI and sex differences. Finally - back to the beginning, the question of sex differences in moral development. The evidence for such differences in the ECI has been quite complex. Most studies have not found significant differences between males and females in average ECI levels during adolescence or young adulthood (e.g., Pratt *et al.,* 2004; Skoe, 2010; Skoe & Diessner, 1994; Sochting, Skoe & Marcia, 1994). As an example, in the study by Skoe and Lippe (2002), the mean ECI levels for women ($M = 2.20$, $SD = .53$) and for men ($M = 2.18$, $SD = .53$), were both between levels 2 and 2.5, $F(1, 142) = .02$, $p = .883$, *ns*; the magnitude of the effect size was close to zero (Cohen's $d = -.04$). In later adulthood, however, women scored higher than men in two independent Canadian samples (Skoe et al., 1996). Studies among children and early adolescents showed a similar sex difference (favouring girls) in the U.S. (Meyers, 2001) and Canada (Skoe & Gooden, 1993). Skoe et al. (1999) compared care-based moral reasoning measured with the ECI in Norwegian early adolescents with data obtained previously from Canadians of the same age (Skoe & Gooden, 1993). In the Canadian sample, girls ($M = 1.77$, $SD = 0.26$) scored higher (near the mid-point of level 2) than did boys ($M = 1.51$, $SD = 0.31$), $F(1, 44) = 8.94$, $p < .01$; the magnitude of the effect size was large favoring girls ($d = -.91$). In contrast, Norwegian girls ($M = 1.48$, $SD = 0.26$) scored similar to Norwegian boys ($M = 1.51$, $SD = 0.30$), $F(1,77) = .33$, *ns;* the magnitude of the effect size was trivial favoring boys ($d = .11$). Hence, sex differences

may be bound with culture. A meta-analysis on the ECI (covering seven independent samples) indicated a small to moderate advantage for females ($d = -.34$) that appeared larger among middle-aged and older adults than among adolescents and young adults (Jaffee & Hyde, 2000). Moderator analyses for age differences could not be conducted, however, due to the small sample of studies.

Nonetheless, the ECI was found to be more strongly related to identity development for women than for men as noted above (Skoe & Diessner, 1994) and to androgynous gender role orientation for women only (Skoe, 1995; Sochting et al., 1994). These findings suggest that care reasoning development is more central to women's than to men's personality development. So, although research results with the ECI suggest that sex differences are not as pronounced as claimed by Gilligan, they also indicate that such differences may be more subtle and complex than simple main effects on standardized measures. Several samples were restricted to university student populations, and sex differences (perhaps cultural differences as well) may be minimized in such samples. Many additional variables must be considered in understanding moral development besides sex, such as cohort variations, gender roles, situational context, sample and location characteristics, stage or period in life, cultural background and religious experience.

ECI and the Model of Hierarchical Complexity. The levels of the care ethic elicited in the ECI differ in their intrinsic level of complexity in how the notion of care is conceived. Therefore, it is of direct relevance to the Model of Hierarchical Complexity (MHC), which this special edition is focused on. In order to provide a framework for studying the links between these two models, a suggested set of parallels between ECI levels and the MHC stages is provided in Table 1. This was devised by establishing parallel degrees of complexity across the levels and stages, in conjunction with the originator of MHC (M. L. Commons, personal communication, December 12, 2010). Each of the ECI levels can be linked to the gradual growth of a capacity to reason about a system of variables (self and other's desires, needs, and welfare) in increasingly complex ways, as shown in the table. It would be interesting to study how the MHC stages might serve as possible enabling capacities in the development of care reasoning over the life course. At the *systematic stage*, a person can solve problems that have multiple causes and/or multiple solutions; it is more complex than formal operational thinking, which tends to conceive single causes and solutions (Robinson, 2012). It seems reasonable to expect a more integrative cognitive ability to aid problemsolving in the relational domain.

Within the Commons' system, higher stages beyond systematic capabilities are also described. This raises the interesting possibility that we might consider the idea of a higher ECI level that takes a wider, societal perspective, involving care also for something beyond or larger than the interpersonal relationship itself. One could imagine, for example, an ECI level 3.5, balanced care for self and other in interaction with society, which may correspond with MHC stage 13 of *Metasystematic* reasoning. Using the Commons' framework to think about this, both self/other relations and society could be viewed as systems that must be brought into interaction and mutual coordination in the person's thinking. The societal level, in which the self/other relationship is embedded, includes a

sense of care and responsibility to society, the environment and future generations; this level may, in turn, be a transition to a hypothetical ECI level 4 that takes a broader, more universal or even cosmic perspective on the transactions between self, other, the wider world, and life itself.

The parallels suggested in Table 1 can be validated empirically, using both cross-sectional and longitudinal methods. Cross-sectionally, MHC stage and ECI level could be correlated in differing groups and cultures to explore the proposed linkages. Longitudinally, transitions between the levels and stages should be closely linked in terms of chronological timing, if the proposed equivalency in Table 1 is correct. These predictions provide a basis for an avenue of integrative research that makes links between the two theories and their measures. In addition, there is a certain theoretical "space" occupied by a number of concepts having to do with the development of progressively more complex and inclusive modes of thinking. This is after all what Piaget had in mind when he spoke of "genetic epistemology". Occupants of this space would include, among others, ego development (Loevinger), ego identity (Erikson, Marcia), care-based moral thought (Gilligan, Skoe), justice-based moral thought (Kohlberg), "mentalization" (Fonagy, Gergely, Jurist, & Target, 2002), "the evolving self" (Kegan, 1982) and "hierarchical complexity" (Commons, 2008). Identifying the common characteristics of these approaches and their underlying cognitive, affective and interpersonal roots could furnish a unified theory that would give us a predictive basis for understanding lifespan development.

Retrospect and prospect

Overall, then, the ECI appears to be a stable and sensitive test of an individual's developmental level of care-based moral thought. This instrument generally has good internal consistency and yields reliable scores even with self-trained raters. More than 1,500 participants have so far been assessed, of both sexes, of several nationalities, and ranging in age from 10 to 85 years of age. Construct validation of the ECI has been an ongoing process closely involved with theory development and research results (Gibbs & Widaman, 1982).

The ECI is a unique measure of care-reasoning in that it assesses developmental care levels, can be used in a wide age range with both men and women, takes only about 30 minutes to administer and score, does not require any external or extensive self-training (the manual is only 26 pages long), and yields substantial and reliable information about the participant's personal and interpersonal or psychosocial functioning. Furthermore, it can be used in several countries with different languages without much translation/back-translation work, as only the dilemmas need to be translated. Both interviewers and participants usually find the interview both interesting and enjoyable. In short, the ECI is a relatively inexpensive and efficient tool for assessing an individual's level of care-based moral thought.

Table 1. A suggested relationship between Skoe's ethic of care interview and Common's model of hierarchical complexity

The ethic of care interview		Model of hierarchical complexity	
Level		**Stage**	
Number	**Name**	**Number**	**Name**
1.0	Survival (caring for self)	8	Primary
1.5	Transition from self-care to responsibility	9	Concrete
2.0	Self-sacrifice (caring for others)	10	Abstract
2.5	Transition to reflective care	11	Formal
3.0	Balanced care for self and others	12	Systematic
3.5	Integrated care for self and others	13	Metasystematic

A series of studies has shown that variations in care-reasoning levels have implications for personal and interpersonal adaptation across the lifespan. The ECI levels are strongly and positively related to identity and personality development, as well as to pro-social interactions and behavior. They are also linked to cognitive abilities such as perspective-taking, justice reasoning and verbal intelligence. Individuals higher in care development appear to have a greater ability to cope with conflicts and other people's distress, to tolerate ambiguity, and to balance their concern for others and self. The highest levels of ethic of care represent the integration of capacities for autonomy and for intimacy (Skoe, 2008; Skoe & Lippe, 2002). Research has shown, however, that in normal groups people rarely score at the highest care level; generally, only around 15% are rated at ECI level 3 (Skoe, 1998). The question then arises: What kind of mechanisms or factors may promote change and growth in care-based moral thought?

Although there has been less research on care than on justice aspects of moral thought, there is a growing body of work on pro-social concerns in moral conflicts (e.g., Carlo, 2006; Eisenberg *et al.*, 2005) and on care-based moral development (e.g., Juujärvi, 2005, 2006; Juujärvi et al., 2010; Skoe, 2010, 2013). Most of this research has been conducted on children, adolescents and young academic populations. Research on moral thought and understanding in adults remains surprisingly scarce. As people grow older, they likely go through experiences that initiate thinking and re-evaluation of life, moral values, self and relationships. In adulthood people usually encounter such life issues as establishing long-term relationships, marriage, home, children, serious career decisions, leadership or work responsibilities, taking care both of one's children as well as aging parents, and coping with one's own aging process. These challenges of maturity, all of which involve care for both oneself and others, may help to move persons forward toward the higher ECI levels.

In addition to the effects of such normative challenges, it has been argued that crisis reveals, as well as creates, character. Encounter with stress and conflict provide an opportunity for positive development, although it may also lead to moral nihilism (Gilligan, 1982). If this is true, one might predict that certain types of crises or painful events will stimulate personal growth, for example, turbulent divorce or breakups of important relationships, or serious illness or death of significant others. Considering the use of care-based moral thought in the broader context of life span

development, more sophisticated capacities in reasoning about care of self and others may serve as an important psychological resource in adapting to the central tasks of adulthood, such as achieving intimacy, a sense of generativity and ego integrity (e.g., Erikson, 1982; Skoe et al., 1996). Each one seems to integrate these two components (self and other), although somewhat differently, across adulthood. Hence, balanced care for self and others may be an important part of maturity or wisdom (Skoe, 1998, 2008).

Adulthood is a time when many begin to question the meaning of their life in the long run and in the face of their own mortality. Such existential questioning certainly could be related to care-reasoning development toward higher and more sophisticated levels. For example, Noddings (2002) noted that for people living in a violent world, the search for meaning is especially important; engaging in such a search is a sign of caring for oneself, and part of learning to care for self is "a concomitant learning to care for others" (p. 35). Her words echo those of Kohlberg (1970) who wrote that we must find meaning in our own lives before we can find it in helping others. In his lecture on a metaphoric Stage 7, Kohlberg (1970) said "To answer the question of why be moral is to tell you the meaning of life, to give you faith" (p. 1). Faith may also be an answer to "Why care?" Perhaps finding meaning in life is a key to gaining insight, not only into human interconnection, but also into one's unity with the cosmos, nature or God (Kohlberg & Ryncarz, 1990). ■

REFERENCES

Adshead, G., Brown, C., Skoe, E. E., Glover, J., & Nicholson, S. (2008). Studying moral reasoning in forensic psychiatric patients. In G. Widdershoven, J. McMillan, T. Hope and L. van der Scheer (Eds.), *Empirical ethics in psychiatry* (pp. 211-230). New York: Oxford University Press.

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*(6), 1173-1182. doi: 10.1037/0022-3514.51.6.1173

Blasi, A. (1998). Loevinger's theory of ego development and its relationship to the cognitive-developmental approach. In P. Michiel Westenberg, A. Blasi, & L. D. Cohn (Eds.), *Personality development: Theoretical, empirical, and clinical investigations of Loevinger's'conception of ego development* (pp. 13-25). London: Erlbaum.

Carlo, G. (2006). Care-based and altruistically based morality. In M. Killen & J. Smetana (Eds.), *Handbook of moral development* (pp. 551-579). Mahwah, New Jersey: Erlbaum.

Colby, A., & Kohlberg, L. (Eds.), (1987). *The measurement of moral judgment.* New York: Cambridge University Press.

Commons, M. L. (2008). Introduction to the model of hierarchical complexity and its relationship to postformal action. *World Futures, 64*(5-7), 305-320. doi:10.1080/02604020802301105

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*(1), 113-126. doi:10.1037/0022-3514.44.1.113

Davis, M. H. (1996). *Empathy - A social psychological approach.* Boulder, CO: Westview.

Eisenberg, N., Cumberland, A., Guthrie, I. K., Murphy, B. C., & Shepard, S. A. (2005). Age changes in prosocial responding and moral reasoning in adolescence and early adulthood. *Journal of Research on Adolescence, 15*(3), 235-260. doi:10.1111/j.1532-7795.2005.00095.x

Erikson, E. H. (1982). *The life cycle completed. A review.* New York: Norton.

Fonagy, P., Gergely, G., Jurist, E., & Target, M. (2002). *Affect regulation, mentalization, and the development of the self.* New York: Other Press.

Gibbs, J. E., & Widaman, K. F. (1982). *Social intelligence: Measuring the development of sociomoral reflection.* Englewood Cliffs: Prentice-Hall.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development.* Cambridge, MA: Harvard University Press.

Haan, N. (1975). Hypothetical and actual moral reasoning in a situation of civil disobedience. *Journal of Personality and Social Psychology, 32*(2), 255-270. doi:10.1037/0022-3514.32.2.255

Hauser, S. T. (1976). Loevinger's model and measure of ego development: A critical review. *Psychological Bulletin, 83*(5), 928-955. doi:10.1037/0033-2909.83.5.928

Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice.* New York: Cambridge University Press.

Hy, L. X., & Loevinger, J. (1996). *Measuring ego development.* New Jersey: Erlbaum.

Jaffee, S., & Hyde, J. S. (2000). Gender differences in moral orientation: A meta-analysis. *Psychological Bulletin, 126*(5), 703-726.

Juujärvi, S. (2005). Care and Justice in Real-Life Moral Reasoning. *Journal Of Adult Development, 12*(4), 199-210. doi:10.1007/s10804-005-7088-7

Juujärvi, S. (2006). The ethic of care development. A longitudinal study of moral reasoning among practical-nursing, social-work and law-enforcement students. *Scandinavian Journal of Psychology, 47*(3), 193-202. doi:10.1111/j.1467-9450.2006.00507.x

Juujärvi, S., Myyry, L., & Pesso, K. (2010). Does care reasoning make a difference? Relations between care, justice and dispositional empathy. *Journal of Moral Education, 39(4)*, 469-489. doi:10.1080 /03057240.2010.521381

Kegan, R. (1982). *The evolving self: problem and process in human development.* Cambrige, MA: Harvard University Press.

Kohlberg, L. (December, 1970). *The Ethical Life, the Contemplative Life and Ultimate Religion − Notes Toward Stage 7 − Social Sciences 154 lecture.* Unpublished manuscript, Pusey Archives, Harvard University. (HUGFP 130.54, Box 3, "Notes Toward Stage 7").

Kohlberg, L. (1973). Continuities in childhood and adult moral development revisited. In P. Baltes & E. W. Schaie (Eds.), *Life-span developmental psychology: Personality and socialization* (pp. 179-204). New York: Academic Press.

Kohlberg, L. (1984). *The psychology of moral development.* San Francisco: Harper & Row.

Kohlberg, L., & Gilligan, C. (1972). The adolescent as a philospher: the discovery of the self in a postconventional world. In J. Kagan & R Coles (Eds.), *12 to 16: Early adolescence* (pp. 144-180). New York: Norton.

Kohlberg, L. & Ryncarz, R. A. (1990). Beyond Justice Reasoning: Moral Development and consideration of a Seventh Stage. In C. N. Alexander & E. J. Langer (Eds.), *Higher Stages of Human Development* (pp. 191-207). New York: Oxford University Press.

Loevinger, J. (1979). Construct validity of the sentence completion test of ego development. *Applied Psychological Measurement, 3*(3), 281-311. doi:10.1177/014662167900300301

Loevinger, J., & Wessler, R. (1970). *Measuring ego development: construction and use of a sentence completion test, Vol. 1.* San Francisco: Jossey-Bass.

Marcia, J. E. (1980). Identity in adolescence. In J. Adelson (Ed.), *Handbook of adolescent psychology* (pp. 159-187). New York: Wiley.

Marcia, J. E., Waterman, A. S., Matteson, D. R., Archer, S. L., & Orlofsky, J. L. (1993). Ego *identity: A handbook for psychosocial research*. New York: Springer−Verlag.

Meyers, C. S. (2001, October). *Care-based reasoning of early adolescents: Its relationship with sex differences, feminine gender role identity, and teacher-perceived cooperation*. Paper presented at the Association for Moral Education Conference. Vancouver, B.C. Canada.

Noddings, N. (2002). *Educating moral people. A caring alternative to character education*. New York: Teachers College Press.

Pratt, M. W., Skoe, E. E., & Arnold, M. L. (2004). Care reasoning development and family socialization patterns in later adolescence: A longitudinal analysis. *International Journal of Behavioral Development, 28*(2), 139-147.

Rempel, J. K., Holmes, J. G., & Zanna, M. (1985). Trust in close relationships. *Journal of Personality and Social Psychology, 49*(1), 95-112. doi:10.1037/0022-3514.49.1.95

Rest, J. R. (1979). *Development in judging moral issues.* Minneapolis. University of Minnesota.

Robinson, O. C. (2012). Development through adulthood: An integrative sourcebook. Basingstoke: Palgrave Macmillan.

Sanders, C. E., Lubinski, D., & Benbow, C. P. (1995). Does the Defining Issues Test measure psychological phenomena distinct from verbal ability?: An examination of Lykken's Query. *Journal of Personality and Social Psychology, 69*(3), 498-504. doi: 10.1037/0022-3514.69.3.498

Skoe, E. E. (1986). The development and partial validation of a care-based measure of moral development. (Doctoral dissertation, Simon Fraser University, Burnaby, British Columbia). *Dissertation Abstracts International, 50, 04-B*, p. 1657.

Skoe, E. E. (1993). *The ethic of care interview manual*. Unpublished manuscript available from the author upon request. University of Oslo, Oslo, Norway.

Skoe, E. E. (1995). Sex role orientation and its relationship to the development of identity and moral thought. *Scandinavian Journal of Psychology, 36*(3)*,* 235-245.

Skoe, E. E. A. (1998). The ethic of care: Issues in moral development. In E. E. Aspaas Skoe & Lippe, A. L. von der (Eds.) *Personality development in adolescence: A cross national and life span perspective* (pp. 143-171). London: Routledge.

Skoe, E. E. (2008). Care, Inventory of (Ethic of Care Interview). In F. Clark Power, R. J. Nuzzi, D. Narvaez, D. Lapsley, & T. C. Hunt (Eds.), *Moral Education. A Handbook. Vol. 1: A-L* (pp. 57-58). Westport, CT: Praeger.

Skoe, E. E. A. (2010). The relationship between empathy-related constructs and care-based moral development in young adulthood. *Journal of Moral Education, 39*(2), 191-211. doi: 10.1080/03057241003754930

Skoe, E. E. A. (2013). The Ethic of Care: Theory and Research. In B. J. Irby, G. Brown, R. Lara-Alecio, & S. Jackson (Eds.), *The Handbook of Educational Theories* (pp. 614-629*).* Information Age Publishing Inc.

Skoe, E. E. A., Cumberland, A., Eisenberg, N., Hansen, K., & Perry, J. (2002). The influences of sex and gender-role identity on moral cognition and prosocial personality traits. *Sex Roles, 46*(9-10), 295-309. doi:10.1023/A:1020224512888

Skoe, E. E., & Diessner, R. (1994). Ethic of care, justice, identity and gender: An extension and replication. *Merrill-Palmer Quarterly, 40*(2), 102-119.

Skoe, E. E. A., Eisenberg, N., & Cumberland, A. (2002). The role of reported emotion in real-life and hypothetical moral dilemmas. *Personality and Social Psychology Bulletin, 28*(7), 962-973. doi:10.1177/01467202028007009

Skoe, E. E. & Gooden, A. (1993). Ethic of care and real-life moral dilemma content in early adolescents. *Journal of Early Adolescence*, *13*(2), 154-167. doi:10.1177/0272431693013002002

Skoe, E. E. A, Hansen, K. L., Mørch, W. T., Bakke, I., Hoffmann, T., Larsen, B., & Aasheim, M. (1999). Care-based moral reasoning in Norwegian and Canadian early adolescents: A cross national comparison. *Journal of Early Adolescence*, *19*(2), 280-291. doi:10.1177/0272431699019002007

Skoe, E. E., & Lippe, A. L. von der (2002). Ego development and the ethics of care and justice: The relations among them revisited. *Journal of Personality, 70*(4), 485-507.

Skoe, E. E. & Marcia, J. E. (1991). A care-based measure of morality and its relation to ego identity. *Merrill-Palmer Quarterly, 37*(2), 289-304.

Skoe, E. E., Pratt, M. W., Matthews, M., & Curror, S. E. (1996). The Ethic of Care: Stability over time, gender differences and correlates in mid to late adulthood. *Psychology and Aging, 11*(2), 280-292. doi:10.1037/0882-7974.11.2.280

Skoe, E. E. A., Pratt, M. W., & Øvregård, A. (2011, July). *The connections among care-based moral development and trust and commitment in romantic relationships in young adulthood*. Poster presented at the XII European Congress of Psychology, Istanbul, Turkey.

Snarey, J. (1998). Ego development and the ethical voices of justice and care: An Eriksonian interpretation. In P. M. Westenberg, A. Blasi, & L. Cohn (Eds.), Personality development: Theoretical, empirical, and clinical investigations (pp. 163-180). London: Erlbaum.

Sochting, I., Skoe, E. E., & Marcia, J. E. (1994). Care-oriented moral reasoning and prosocial behavior: A question of gender or sex role orientation. *Sex Roles, 31* (3-4), 131-147. doi:10.1007/BF01547711

Walker, L. J. (2006). Gender and morality. In M. Killen & J. Smetana (Eds.) Handbook of moral development (pp. 93-115). Mahwah, New Jersey: Erlbaum.

Wark, G. R., & Krebs, D. L. (1997). Sources of variation in moral judgment: Toward a model of real-life morality. *Journal of Adult Development*, *4*(3), 163-178.

APPENDIX

» THE ETHIC OF CARE INTERVIEW

The Researcher Generated Dilemmas

The specific researcher generated dilemmas for *females* are as follows:

The Lisa dilemma

Lisa is a successful teacher in her late twenties who has always supported herself. Her life has been centered on her work and she has been offered a permanent position for next year. Recently she has been involved in an intense love affair with a married man and now finds that she is pregnant.

What do you think Lisa should do? Why?

The Betty dilemma

Betty, in her late thirties, has been married to Erik for several years. They have two children, 8 and 10 years old. Throughout the marriage Betty has been at home, looking after the house and the children. For the last few years Betty has felt increasingly unhappy in the marriage relationship. She finds her husband demanding, self-centered and insensitive as well as uninterested in her needs and feelings. Betty has several times tried to communicate her unhappiness and frustration to her husband, but he continually ignores and rejects her attempts. Betty has become very attracted to another man, Steven, a single teacher. Recently, Steven has asked Betty for a more intimate, committed relationship.

What do you think Betty should do? Why?

The Kristine dilemma

Kristine, a 26-year-old woman, has decided to live on her own after having shared an apartment with a girlfriend for the last three years. She finds that she is much happier living alone as she now has more privacy and independence and gets more work and studying done. One day her mother, whom she has not seen for a long while as they do not get along too well, arrives at the doorstop with two large suitcases, saying that she is lonely and wants to live with Kristine.

What do you think Kristine should do? Why?

The specific researcher generated dilemmas for *males* are as follows:

The Derek dilemma

Derek is a married, successful teacher in his late twenties. His life has been centered on his work and he has been offered a permanent position for next year. Recently, he has been involved in an intense love affair with a single woman who has just told him that she is pregnant and that it is his child.

What do you think Derek should do? Why?

The Erik dilemma

Erik, in his late thirties, has been married to Betty for several years. They have two children, 8 and 10 years old. Throughout the marriage Betty has been at home, looking after the house and the children. For the last few years Erik has felt increasingly unhappy in the marriage relationship. He finds his wife demanding, self-centered and insensitive as well as uninterested in his needs and feelings. Erik has several times tried to communicate his unhappiness and frustration to his wife, but she continually ignores and rejects his attempts. Erik has become very attracted to another woman, Carol, a single teacher. Recently, Carol has asked Erik for a more intimate, committed relationship.

What do you think Erik should do? Why?

The Chris dilemma

Chris, a 26-year-old man, has decided to live on his own after having shared an apartment with a friend for the last three years. He finds that he is much happier living alone as he now has more privacy and independence and gets more work and studying done. One day his father, whom he has not seen for a long while as they do not get along too well, arrives at the doorstop with two large suitcases, saying that he is lonely and wants to live with Chris.

What do you think Chris should do? Why?

**BEHAVIORAL
DEVELOPMENT
BULLETIN**

Relationship among measures within the social and moral development domain

Eva Yujia Li[1], Michael Lamport Commons[2], Jonas Gensaku Miller[3], Terri Lee Robbinet[4], Helena Marchand[5], Carrie Melissa Ost[6], and Sara Nora Ross[7]

1 Harvard Graduate School of Education
2 Harvard Medical School
3 University of California at Davis
4 University of Phoenix
5 University of Lisbon
6 Dare Institute
7 Antioch University Midwest McGregor

ABSTRACT

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.

KEYWORDS: moral developmental, model of hierarchical complexity, developmental stage, instruments to measure stage, preference

In the modern era of multi-cultural societies, international relations, and the "War on Terror", it is more important than ever to understand how clashes between moral belief systems can be reconciled. In order to understand the interaction between belief systems, however, we must attain a better understanding of how moral reasoning develops in each individual. Just as each individual passes through stages of other forms of development, a person's performance in reasoning about moral issues develops in a series of stages.

In this study, we investigate moral development as it relates to three very complicated and controversial issues. We designed each of these dilemmas to strike at the core of participants' moral sense. This should make it more likely for each participant to seriously consider and reflect on the reasoning that backs their often firmly held beliefs. The issues we investigated were: whether or not to report incestuous rape; the acceptability of capital punishment,

and informed consent between a counselor and a patient. In investigating these three topics, we hope to determine the stages of moral development with which people reason about these issues and how well the Model of Hierarchical Complexity accounts for their performance. We also expect to compare how groups of individuals reasoned on each of the three issues.

Theories of moral development

Piaget's studies of moral judgment can be summarized by a two-stage theory, with a transition of the form of moral reasoning occurring typically around age 10 and 11. Consistent with Piaget's notion of what consists of a development stage, Kohlberg outlined six stages of moral development, including two stages at the Pre-conventional level, two at the Conventional level, and two at the Post-conventional level. These stages describe the cognitive development of moral reasoning and have been the mainstream of moral development studies. He explained that development happens through socialization and thinking about moral issues, not through unfolding of genetic blueprints (Crain, 1985).

Author note: Correspondence regarding this article should be directed to Michael Lamport Commons, Harvard Medical School, 234 Huron Avenue, Cambridge, MA 02138. E-mail: commons@tiac.net

Table 1. Orders of hierarchical complexity

Order	Name complexity	Order	Name complexity
0	Calculatory	8	Concrete
1	Sensory & motor	9	Abstract
2	Circular sensory-motor	10	Formal
3	Sensory-motor	11	Systematic
4	Nominal	12	Metasystematic
5	Sentential	13	Paradigmatic
6	Preoperational	14	Crossparadigmatic
7	Primary	15	Metacrossparadigmatic

Note. MHC stages and order numbers have not been revised in this version.

Kohlberg's model of moral development emphasized the structural differences between each stage of development: Every two stages have qualitative differences; every stage is a structured whole; there is an invariant sequence of development. A higher stage is a hierarchical integration of lower stages; and there is universal sequence. The most notable criterion is the hierarchical structure of the model, which stated that the higher stages were integration of lower stages. This depicted the structure of the model and explained the sequence (Crain, 1985).

There have been several streams of criticism on Kohlberg's model of moral development. Some authors question the cross-cultural validity of the model of moral developmental stages, arguing while the trend toward maturity is universal, high stages in Kohlberg's model may be culturally specific (Gibbs, Basinger, Grime & Snarey, 2007).

Some argued that there are gender differences in moral developmental processes, as women are socialized in a different way from men. Caring and Justice may be somewhat different domains (Bill, 1994). Robbinet (2008) argued that the methods of measuring moral development may be biased, as past studies found high correlations between high moral stages and liberal political ideology. This shows a possible political bias in the Kohlberg scoring of dilemmas. Robbinet (2008) used carefully constructed moral vignettes to assess moral development and did not find any relationship to political affliction.

The arguments call into question the potential for content bias of moral developmental stages. This paper is an effort to provide more objective and less content dependent measures to assess the stages of moral development. The vignettes to assess moral development were based on Model of Hierarchical Complexity, an analytical framework which focuses on the structure of moral reasoning. Items were constructed to address moral reasoning at different stages. Rasch Analysis was used to check that the item difficulties were consistent with their subjective stage of moral reasoning. Three instruments were constructed using the model, all of which have the same structure and some variation in content form.

Another potential shortcoming of both Piaget and Kohlberg's models was that they were not detailed enough to capture development accurately. Kohlberg's stages started at some indeterminate point in childhood, rather than starting from birth. The Model of Hierarchical Complexity addresses developmental stages starting from birth and throughout the lifespan.

The model of hierarchical complexity: A developmental stage theory

The Model of Hierarchical Complexity (Commons & Pekker, 2008; Commons, Trudeau, Stein, Richards, & Krause, 1998) forms the basis for a stage theory that is applied to explain development in multiple domains. According to the theory, development can be measured by the hierarchical complexity of tasks that an individual successfully addresses. The complexity of tasks is measured by applying three axioms.

1. First, a more hierarchically complex task is defined in terms of two or less hierarchically complex ones from the next order below.

2. Second, the more hierarchically complex task organizes or coordinates two or more less complex ones. That is, the more complex task specifies the way in which the less complex ones combine.

3. Third, the coordination of tasks that occurs must be non-arbitrary. Figure.1 illustrates the relationship between higher order tasks and lower order tasks. Past research has identified 16 orders of Hierarchical Complexity, as shown in Table 1. The person's performance in completing the task is called the Stage of Performance. For example, reasoning about complex moral issues is a task. The Order of Hierarchical Complexity at which the individual reasons about the moral issue reflects the person's Stage of Performance on that task.

What sets the Model of Hierarchical Complexity (Commons & Pekker, 2008; Commons, Trudeau, Stein, Richards, & Krause, 1998) apart from other theories of moral development is its focus on the general structure of development rather than development within a specific domain. For example, it can be applied to studying development of math and science knowledge, social perspective taking, etc. Tasks of different domains can be scored (Commons, Danaher – Gilpin, Miller & Goodheart 2002). The core of the model is the idea that as development occurs, individuals become increasingly able to accomplish complex tasks that coordinate and are defined by lower order tasks. Therefore, it does not depend on content, culture or context in formulating stages. When the model is applied to a specific domain, stages of development have to be generated within the domain, using the three axioms of tasks stated above. Many empirical studies have been done to test the validity of the developmental stages generalized by the Model of Hierarchical Complexity (Commons, Goodheart, Dawson, Draney, Adams & Marie, 2008; Commons, Rodriguez, Adams, Goodheart, Thomas & Ellen, 2006). In particular, analytic work has been done to explain the relationship between the Model of Hierarchical Complexity and Kohlberg Moral Development Stages.

Using model of hierarchal complexity to build a suggested sequence of moral developmental stages

The Model of Hierarchical Complexity will be used initially in this paper to generate a proposed complete sequence of moral developmental stages, including the precursors to Kohlberg's stages of moral development. To do this, we began with the most elemental actions and perceptions, at Stage 1. We then build them up stage by stage. Each new stage's actions are defined in terms of the lower stage actions, and the actions organize themselves in a non-arbitrary way. The stages below have not incorporated the new revisions made to the MHC stages and order numbers.

Stage 1. Sensory or Motor: Infants have basic feelings of pain and pleasure. There may be early expressions of reflexive empathy, which is seen in that young infants will cry when other infants are also crying. Infants can either actively perceive things or emotionally or motorically act. However, they cannot coordinate the two except in a reflexive way.

Stage 2. Circular Sensory Motor: Infants begin to coordinate their parent's perceived emotions with their own behaviors. They look for preferred caregivers and reach for them. They may protest over loss and show joy over such things as reunion. Through interactions, infants share pleasure with caregivers, which can form a basis for caring. Caring for others is necessary of developing empathy later on, which is necessary but not sufficient action for moral judgment and action. Operant imitation, or imitative behavior that is a function of its consequence, develops. For example, an infant may imitate a parent's speech. This takes the form of babbling in which the phonemes match the phonemes in the parents' speech. The babbling may be reinforced as the parent smiles at the child's babbling. This forms the basis of modeling and identification of moral behaviors later on.

Stage 3. Sensory-Motor: Infants develop vague concept of right or wrong. The emotions shown by their attachment figures serves as a reinforcer for the behavior that pleases those adults. They show understanding of fairness and prefer equal distribution of resources to unequal distribution (Sloane, Baillargeon & Premack, 2012; Geraci & Surian, 2011). Infants displays consoling type (or empathic) responses when someone else is upset. These responses involve only the infant's own body. Pats another person, hugs them, or looks concerned.

Stage 4. Nominal stage: Infants understand "no." They may slowly push something to the edge of table and watch the mother's reaction. They show pride when they do things right without requiring approval. They demonstrate that they may inhibit behavior. They reflect on "greater than" and "less than" from the previous stage. They may reflect upon whether or not a portion was fair. This is necessary to see injustice. They recognize self-versus-other but cannot make comparisons.

Stage 5. Sentential (Stage 1 Kohlberg): Toddler says "I good". They form simple sentences by putting two words in an order together. One is "I" and the other is "good" yielding "I good." Each word is from the nominal stage. They also develop a sense of shame.

Stage 6. Preoperational (Stage 1 / 2 Kohlberg): Children coordinate multiple sentences and are may tell a whole story about good or bad – such as a part of a fairy tale, I was bad and am now good.

Stage 7. Primary (Stage 2 Kohlberg): Children take their own perspectives. They know what they like and value. They look for "what's in it for me". They follow rules to avoid punishment. They understand power relations. They know that the rules of authorities should be followed and they may attempt to impose rules on others. They also know what someone else likes and values, but they do not yet coordinate their own perspective and those of others.

Stage 8. Concrete (Stage 2/3 Kohlberg): Individuals take the perspective of another and integrate it with their own perspectives. By integrating both perspectives, individuals make fair deals between the two people. However, fairness is understood only among two or a few people. Individuals obey authority.
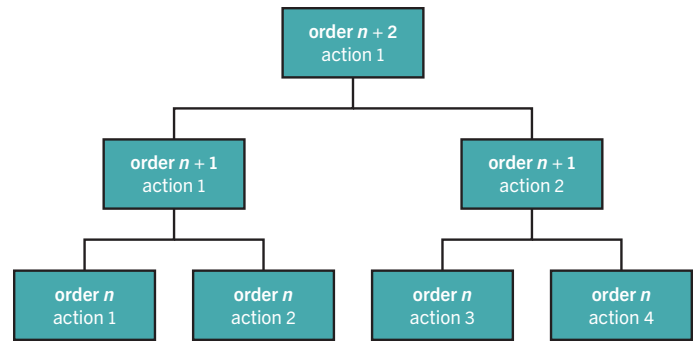

Figure 1. Order of hierarchical complexity of tasks

Stage 9. Abstract (Stage 3, Kohlberg): Children take perspective of a group. They understand social norms, such as what is being "good' or "bad". Quantification words like "everyone in my group" appear. Children may reason about what others think. Children understand personalities, traits and other variables. The dimensionalized qualities may be used to express preferences.

Stage 10. Formal (Stage 3/4 Kohlberg): Discussions are logical and empirical support is logically brought. Words like "if ...then," "in every case, it turned out the same," "the reasons were" occur. This is the stage with univariate and linear explanations. There can be multiple outcomes however. The different outcomes are generally unrelated so they do not form systems.

Stage 11. Systematic (Stage 4, Kohlberg): The simple linear relationships from formal operations are inter-coordinated into systems. Words like bureaucratic, capitalist, functional, and structural that describe systems of relationships appear. The logical structure of this stage coordinates multiple aspects of two or more abstractions, as in: "relationships are built on trust and though we cannot always keep them, making promises is one way we build trust, so it's generally better to make promises than not to make them." Here, the importance of trust to relationships, building trust, and the possibility that promises can be broken, are all taken into account while formulating the conclusion that promises are desirable. Each system consists of multivariate inputs or multiple relations. For example, A or B causes C can be decomposed into two causal relations, A causes C or B causes C. A and B causes C is the cross product of two independent variables. Think of systems as a two or more way ANOVA or a regression equation with cross products and multiple inputs.

Stage 12. Metasystematic (Stage 5, Kohlberg): the new concepts are referred to as 1st order principles. These coordinate formal systems. Words like autonomy, parallelism, heteronomy, and proportionality are common. The metasystematic stage concept of parallelism, for example, can be employed to compare the structures of the military and of camp as institutions. The logical structure of this stage identifies one aspect of a principle or an axiom that coordinates several systems, as in: "contracts and promises are articulations of a unique human quality, mutual trust, which coordinates human relations." Here, contracts and promises are seen as the instantiation of a broader principle coordinating human interactions.

Table 2. General description of sequence

MHC	Kohlberg & descendants	Discriminations
0	-1	Calculatory
1	0/-1	Sensory & motor actions
2	0	Circular sensory-motor actions
3	0/1	Sensory-motor
4	—	Nominal
5	1	Sentential
6	1 /2	Preoperational
7	2	Primary
8	2/3	Concrete
9	3	Abstract
10	3/4	Formal
11	4	Systematic
12	5	Meta-systematic
13	6	Paradigmatic
14	7	Cross-paradigmatic

Note. MHC stages and order numbers have not been revised in this version.

Few individuals perform at stages above formal operations. Only 20% of the people perform at Systematic stage and 1.5% at Metasystematic stage (Kallio, 1995; Kallio & Helkama, 1991). Some adults are said to develop alternative to, and perspectives on, formal operations. They use formal operations within a "higher" system of operations and transcend the limitations of formal operations. In any case, these are all ways in which these theories argue and present converging evidence that adults are using forms of reasoning that are more complex than formal operations.

The MHC accounts for all of Kohlberg's stages except for Moral stage 4/5 which the model asserts is just a transition between Kohlberg's stage 4 and 5. As with Piaget, following Pascual-Leone (1970), all the half stages are shown to be full stages. Table 2 illustrates the relationship between Moral Developmental Stages of Model of Hierarchical Complexity and Kohlberg's stages of moral development.

Current study

In this study, it will be investigated as to how well the Orders of Hierarchical Complexity accounts for the difficulty of performance on moral reasoning tasks. To better understand how people think about moral issues we constructed a dilemma for each issue, followed by a series of arguments and lines of reasoning that evaluate the issue at hand. Each argument was constructed at a given Order of Hierarchical Complexity. Each participant rated the quality of all of the arguments.

The Order of Hierarchical Complexity should predict the difficulty of lines of reasoning in the instruments. The line of Reasoning with the highest Order of Hierarchical Complexity should be the most difficult, and vice versa. Rasch Analysis will be used to estimate the difficulty of items (1960/1980). If the results of this study support this prediction, this will validate the stages of moral development based on Model of Hierarchical Complexity.

» METHOD

Participants

There were four samples of participants, all of whom were obtained on-line. One sample of 103 participants completed the *Counselor Patient Instrument*, which was sent to various e-mail lists. In the second sample: 96 participants completed *Anti-Death Penalty* Dilemma. In the third sample: 77 participants completed *Incest Rape Dilemma – No Report*. In the last sample: 58 participants completed *Incest Rape Dilemma – Report*. The instrument did not collect demographics information of these participants.

Instrument

Counselor patient instrument. The Counselor Patient Instrument presented five stories, each of which describes how a counselor consults a patient with regard to choosing a treatment to improve the patients' life. In each story, the method in which the counselor consults the patient is of a different Order of Hierarchical Complexity. The methods improved in their coordination of different perspectives. For example: in the concrete order vignette the counselor recommended a method that is recommended by colleagues. The counselor then called in a few colleagues to talk to the patient about the method. In the metasystematic order vignette, counselor explained all aspects of the treatment and describes at length the pros and cons of alternatives, including doing nothing. The counselor asked the patient to consider the discussion they had (the informing system) before making a decision (the consent system).

Participants were asked to a) rate the method of offering the plan of this counselor; b) rate the degree to which this counselor informed their person; c) rate how likely you would be to accept the plan offered by this counselor. Participants answered these questions by rating them the vignettes on a 1 to 6 scale.

Death penalty dilemma. The death penalty dilemma presents five attorneys' arguments *against* capital punishment, ranging from concrete to metasystematic stages. As the order of complexity of the reasoning presented increased, the lines of reasoning improved in their coordination of rights and duties, universality, and the possibility of innocence. For example, the concrete argument simply states, "We do not have the right to take away this person's life," without any support, while the metasystematic argument reasons, "Human rights apply to the worst of us, as well as to the best of us. We are saying that killing another human being is a punishable act, yet we use execution (which is also killing a human being) to condemn murder. Such an act by the government is the mirror image of the criminal's willingness to use physical violence against a victim and should not be condoned."

Participants were asked to a) rate how well each of the attorneys argued the rights and duties of all concerned in this death penalty case; b) rate how likely you are to vote for the death penalty in this case based solely on the arguments of each of the following lawyers; and c) rate how likely you are to vote for mitigation of the death penalty to life in prison with no chance of parole, based solely on the arguments of each of the following lawyers. Participants answered these questions by rating the vignettes on a 1 to 6 scale.

Incest rape dilemma. The incest rape dilemma presents the following problem: "An 18 year old woman has been repeatedly raped by a member of her family since she was 10 years old. The rapist has said that if she reports it, he will rape her sister. Several times she has thought about revealing the situation, but she has not done so yet."

We constructed five lines of reasoning arguing that she should not report the rape, and five lines of reasoning arguing that she should report the rape. Each set of arguments range from concrete 8 to metasystematic stage 12. As the stage of reasoning increases, the lines of reasoning improved in the coordination of rights and duties, concerns of reputation, and the possibility of further harm. For example, the concrete arguments simply argue for one side, while the metasystematic arguments consider the complicated nature of the decision and consider both possibilities before choosing one outcome.

Participants were asked to a) rate each friend's argument, b) rate how well each friend informed the woman, and c) rate how likely you would be to take the advice of the friend. Participants answered these questions by rating the vignettes on a 1 to 6 scale.

Procedure

Each of these instruments was administered online separately.

Data analysis

In this study, we used Rasch Analysis to estimate the difficulty of each item in each vignette. A Rasch Analysis uses probabilistic equations to produce an additive, equal interval scale based on the relationships between how different participants rate particular items on a continuous scale (Rasch, 1980; Wright & Linacre, 2001). Each item on the scale is coded into continuous numeric values (generally between -4 and +4), according to an order of magnitude, which shows the severity of the property of the item. The scale indicates a latent property of items and participants. In the context of this study, items fall on a Rasch scale that indicates difficulty of the items. Rasch scores are called *Rasch Scaled Item Difficulty*.

After analyzing data with a Rasch model, a number of questions can be answered. First, where on the scale does each independent variable fall (e.g. in this case, at what stage is each item). Second, what is the range of scaled values between all variables for all participants? The answer to this question defines the meaning of the difference between scores. For a small range of scaled perceived bias scores, a difference of 1 unit would indicate a big difference, whereas for a large range it would indicate a small difference. Third, what is the scaled value for each participant with regard to hierarchical complexity?

It is also important to examine the extent to which the actual measured items fit the model. This can be determined by infit and outfit MNSQ values, or mean squared residuals (Wright & Linacre, 2001). A large residual indicates a large difference between the model and the actual score. The infit and outfit statistics adopt slightly different techniques for assessing an item's fit to the Rasch model. The infit statistic gives relatively more weight to the performances of persons closer to the item value. The argument is that persons whose ability is close to the item's difficulty should
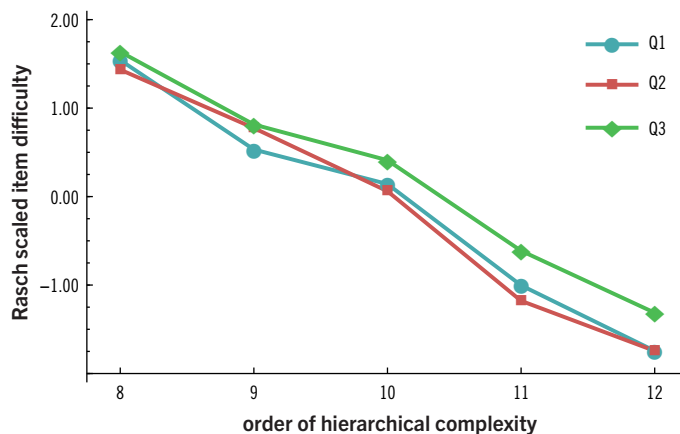

Figure 2. Counselor patient study. The lower (negative) measures indicate a higher rating of difficulty, while higher (positive) measures indicate a lower rating. MHC order numbers have not been revised in this version.

provide a more sensitive insight into the item's performance. The outfit statistic is not weighted, and therefore is more sensitive to the influence of outlying scores. Aberrant infit scores usually cause more concern than large outfit statistics (Bond & Fox, 2001; Linacre, 2002). Nevertheless, both types of scores are presented here. Linacre (personal communication, January, 2003) developed a criterion of rejecting items with infit errors larger than 2.00. He suggested that it is possible that items with an infit score of greater than 2.00 have characteristics that are sensitive to factors not reflected in the scale and may not fit because they are too extreme for the scale or lie on another dimension.

Second, we conducted simple linear regressions of the Rasch Scaled Item Difficulty of items against their Orders of Hierarchical Complexity. In each instrument, there are three questions after each vignette. The items of each question was grouped together and their Rasch Scaled Item Difficulties were used as dependent variables. The independent variables were the Order of Hierarchical Complexity of the same items. This analysis allows us to test whether the Order of Hierarchical Complexity of vignettes predicted their Rasch Scaled Item Difficulties. If the items' Orders of Hierarchical Complexity predicted their Rasch Scaled Item Difficulty, then the result support our theory that moral reasoning fits in the Model of Hierarchical Complexity framework. In the results section, we present the simple correlation between OHC and RSID to show the strength of association, the R^2 of the model to show how much variation in the outcome variable is explained by OHC, and the result of F rest, to show the statistical significance of the regression model.

» RESULTS

Counselor patient

Participants' response to the counselor patient questionnaire showed that the Order of Hierarchical Complexity of the arguments predicted Rasch Scaled Item Difficulty of the arguments. Figure 2 illustrates the relationship between Rasch Scaled Item Difficulty and Order of Hierarchical Complexity. The regression analysis found that the a priori difficulty of the items predicted the participants' ratings of the method of offering the plan to this counselor with a very high r: $r(3) = 0.992$, $F(1, 3) = 188.97$, $p = 0.01$, $r^2 = 0.984$. When
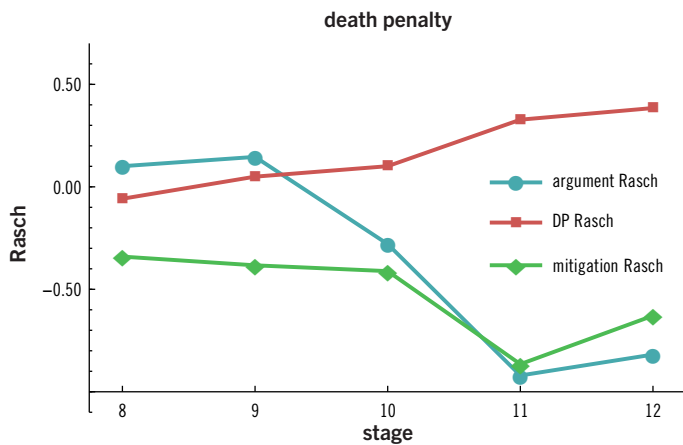
Figure 3. Death penalty dilemma. The lower (negative) measures indicate a higher rating, while higher (positive) measures indicate a lower rating. MHC stages and order numbers have not been revised in this version.

asked to rate the degree to which this counselor informed their person, $r(3) = 0.993$, $F(1, 3) = 203.14$, $p = 0.01$, $r^2 = 0.985$. When asked to rate "how likely you would be to accept the plan offered by this counselor", $r(3) = 0.994$, $F(1, 3) = 259$, $p = 0.01$, $r^2 = 0.989$. The overall regression of the Rasch Scaled Item Difficulty of all items shows that $r(13) = 0.982$, $F(1, 13) = 203.14$, $p = 0.00$, $r^2 = 0.965$.

Anti-death penalty

As shown in Figure 3, when asked to rate how good each argument was, the hierarchical complexity of each line of reasoning predicted the Rasch scaled score, $r(3) = .919$, $F(1,3) = 16.247$, $p = .027$, $r^2 = .844$. Interestingly, the question that yielded the highest correlation to hierarchical complexity in all of the studies was, "How likely are you to vote for the death penalty in this case based solely on the arguments of each of the lawyers?", $r(3) = .921$, $F(1, 3) = 76.226$, $p = .003$, $r^2 = .962$. Yet, the reverse was not true. When asked "How likely are you to vote for mitigation of the death penalty to life in prison with no chance of parole, based solely on the arguments of each of the lawyers?" The correlation between hierarchical complexity and Rasch score, $r(3) = -.764$, $F(1, 3) = 4.199$, $p = .133$, $r^2 = .583$.
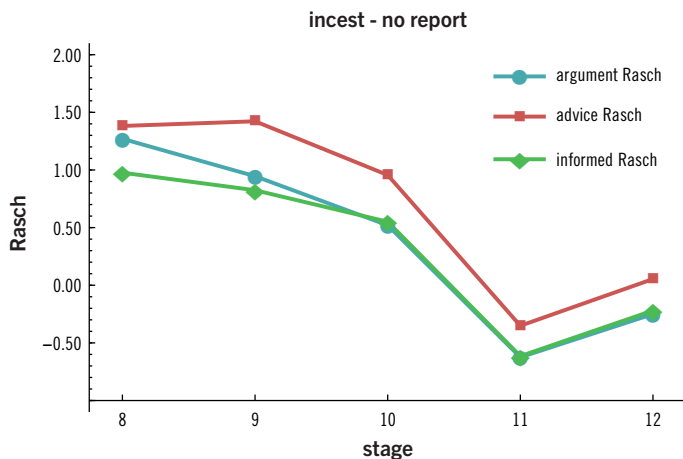

Figure 4. Incest dilemma with result of no report. The lower (negative) measures indicate a higher rating of difficulty, while higher (positive) measures indicate a lower rating. MHC stages and order numbers have not been revised in this version.

Incest—No report

As shown in Figure 4, the Order of Hierarchical Complexity of a line of reasoning strongly predicted its Rasch scaled score in each question. When asked to rate how good each argument not to report was, $r(3) = .916$, $F(1, 3) = 15.676$, $p = .029$, $r^2 = .838$. When asked how well each argument informed the woman, $r(3) = .877$, $F(1, 3) = 10.028$, $p = .051$, $r^2 = .770$. When asked how likely the participant would be to take the advice not to report, $r(3) = -.868$, $F(1,3) = 9.138$, $p = .057$, $r^2 = .753$.

Incest—Report

As shown in Figure 5, the Order of Hierarchical Complexity of a line of reasoning did not predict its Rasch scaled score in each question. When asked to rate how good each argument to report was, $r(3) = .624$, $F(1,3) = 1.918$, $p = .260$, $r^2 = .390$. When asked how well each argument informed the woman, $r(3) = .708$, $F(1, 3) = 3.023$, $p = .180$, $r^2 = .502$. When asked how likely the participant would be to take the advice not to report, $r(3) = .670$, $F(1, 3) = 2.439$, $p = .216$, $r^2 = .448$.

» DISCUSSION

The correlations between Orders of Hierarchical Complexity of the items and the corresponding Rasch scores differed depending on the dilemma. The Counselor Patient instrument yielded the highest predictability. All three questions' Rasch scaled item difficulty were predicted by their Orders of Hierarchical Complexity with r higher than 0.9. This indicates that the Orders of Hierarchical Complexity of the lines of moral reasoning predicted the difficulty of carrying out the task. Therefore, we could say that moral reasoning is a type of task that can be described by the Model of Hierarchical Complexity. The more complex the task, the more difficult it is to carry out. As individuals become capable of accomplishing more and more complex moral reasoning, their moral development occurs.

The death penalty yielded similarly high correlations, particularly on the question that asks the participant whether the argument is compelling enough to sentence someone to death. A flaw in the instrument, however, might have produced the high correlation, however, because the lengths of the arguments varied according to stage. Because the higher stage arguments tended to be longer, participants might have based their choices on length, rather than the vignettes.

When considering the dilemmas together it becomes apparent that variables other than the stage of the argument may have influenced Rasch Scaled Item difficulty somewhat. These differences might have occurred because the nature of the dilemmas and the questions demand different levels of consideration of hierarchical complexity. In the Incest Rape Dilemma and Death Penalty Dilemma, Rasch scores were lower when the question asked if the participant would be willing to act on the line of reasoning, as opposed to when they were just asked to rate the quality of the argument. In other words, participants were more likely to rate vignettes at a higher order of complexity as preferred when the issue being asked about was simply how good the arguments appeared to them. When they were asked to become more in-

volved and were asked about their own actions, they appeared to rate vignettes at a lower level of complexity as more preferred. Alternatively, these discrepancies might suggest that participants were not only basing their choices on the quality of the argument, but brought other factors into their choices such as following low stage beliefs taught to them by their cultures.

In general, the results of incest dilemma, while still showing very high r's, showed less clear patterns than questions pertaining to the death penalty. The Orders of Hierarchical Complexity of arguments in Incest dilemma had less predictability to their Rasch Scaled Item Difficulty than those in Death Penalty Dilemma. This suggests that individuals may think about different situations differently. Although we asked participants to set aside their pre-existing beliefs about these issues, it is likely that these their preconceived notions contributed to how they made their decisions.

The results show that the Order of Hierarchical Complexity accounts for the differences in understanding all of these moral issues. Because the r's are so high, it suggests that order of hierarchical complexity is the major factor that accounts for the moral reasoning shown. So moral action, to the extent to which a choice depends on being at a stage may be determined by three things: a) values as describe by George P. Lakoff (as cited in Robinett, 2006); b) attachment in the classical sense (Ainsworth, Blehar, Waters, & Wall, 1978; Bowlby, 1982); and c) also social perspective taking with attachment stage (Commons, 1991). For those decisions, moral stage is necessary but not sufficient.
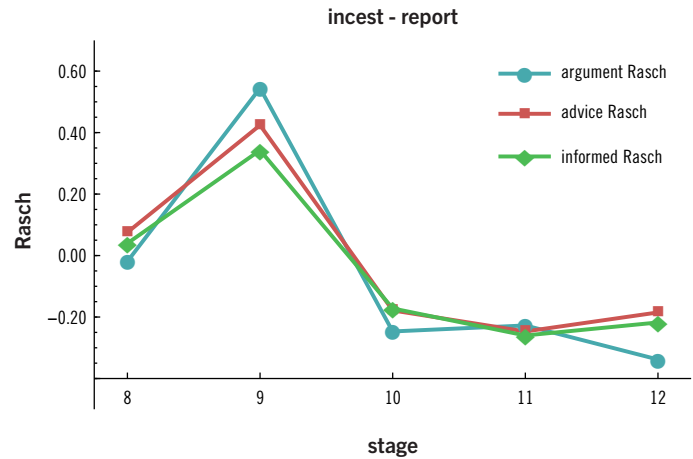

Figure 5. Incest dilemma with result of report. The lower (negative) measures indicate a higher rating, while higher (positive) measures indicate a lower rating. MHC stages and order numbers have not been revised in this version.

To further investigate the correlation between Order of Hierarchical Complexity of tasks and moral reasoning on those tasks, studies similar to this one presenting similar dilemmas might shed light on the trends shown here. Future instruments might try to approach dilemmas that are new to most participants to avoid participants answering according their previous held beliefs. ■

REFERENCES

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment*. Hillsdale, New Jersey: Lawrence Erlbaum.

Bond, T. G., & Fox, C.M. (2001). *Applying the Rasch Model. Fundamental Measurement in the Human Sciences*. New Jersey: Lawrence Erlbaum.

Bowlby, J. (1982). *Attachment*. New York: Basic Books.

Commons, M. L. (1991). *A comparison and synthesis of Kohl-berg's cognitive-developmental and Gewirtz's learning-developmental attachment theories*. In J. L. Gewirtz & W. M. Kurtines (Eds.), *Intersections with attachment* (pp. 257-291). Hillsdale, New Jersey: Lawrence Erlbaum.

Commons, M. L., Danaher-Gilpin, D., Miller, P. M. & Goodheart, E. A. (2002). *Hierarchical complexity scoring system: how to score anything*. Unpublished Scoring Manual.

Commons, M. L., Goodheart, E. P., Dawson, A., Draney, T.L., & 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.

Commons, M.L., Rodriguez, J.A., Adams, K.M., Goodheart, E.A., Thomas, G. C., & Ellen, D. (2006). Informed consent, Do you know it when you see it? Evaluating the adequacy of patient consent and the value of lawsuit. *Psychiatric Annals, 36*(6), 430-435.

Commons, M. L., Trudeau, E. J., Stein, S A., Richards, F.A., & Krause, S. R. (1998). Hierarchical complexity of tasks shows the existence of developmental stages. *Developmental Review. 18*(3): 237-278. doi:10.1006/drev.1998.0467

Crain, W.C. (1985). *Theories of development: Concepts and applications*. New Jersey: Prentice-Hall.

Geraci, A., & Surian, L. (2011). The developmental roots of fairness: Infants' reactions to equal and unequal distributions of resources. *Developmental Science, 14*(5),1012-1020. doi: 0.1111/j.1467-7687.2011.01048.x

Gibbs, J. C., Basinger, K. S., Grime, R. L., & Snarey, J. R. (2007). Moral judgment development across cultures: Revisiting Kohlberg's universality claims. *Developmental Review, 27*(4), 433-500. doi: 10.1016/j.dr.2007.04.001

Kallio, E. (1995). Systematic reasoning: formal or postformal cognition? *Journal of Adult Development*, *2*(3), 187-192. doi:10.1007/BF02265716

Kallio, E., & Helkaman, K (1991). Formal operations and postformal reasoning: A replication. *Scandinavian Journal of Psychology*, *32*(1), 18-21. doi: 10.1111/j.1467-9450.1991.tb00848.x

Linacre J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, *16*(2), 878.

Pascual-Leone, J. (1970). A mathematical model for the decision rule in Piaget's developmental stages. *Acta Psychologica*, *32*, 301-345. doi: 10.1016/0001-6918(70)90108-3

Rasch, G. (1980). *Probabilistic Models for Some Intelligence and Attainment Tests*. Expanded ed. Chicago, Illinois: University of Chicago Press.

Robinett, T. L. (2006). *Moral reasoning and political affiliation in liberal and conservative voters: Applying a model of hierarchical complexity*. Submitted in partial fulfillment of Dissertation, Capella University.

Sloane, S., Baillargeon, R., & Premack, D. (2012). Do infants have a sense of fairness? *Psychological Science*, *23*(2), 196-204. doi: 10.1177/0956797611422072

Wright, B. D., & Linacre, J.M. (2001). Glossary of Rasch measurement terminology. *Rasch Measurement Transactions,15*(2), 824-825.

A pattern recognition method for disclosing different levels of value system from questionnaire data

Per Sjölander[1], Nina Lindström[2], AnnJessica Ericsson[1], and Sofia Kjellström[3]

1 Void Institute, Sweden
2 Southern Lapland Research Department, Sweden
3 Jönköping University, Sweden

ABSTRACT

The aim of the present study was to describe, test and validate a method for disclosing significant response patterns from questionnaire data, and for classifying individual response profiles into a sequence of significant patterns. The method is based on pattern recognition statistics and probability calculations. The results from the population tested show that the method can disclose characteristic profiles of different value systems, and that these systems can be arranged in a hierarchical order similar to the conventional levels of ego development. It is suggested that this method is applicable to any multiple choice-questionnaire containing a number of items where the response alternatives represent a sequential order, for example, of different levels of development within a psychological domain. The method might be a valuable tool for acquiring information on the distribution of different levels of adult development in large populations, such as in communities and large organizations.

KEYWORDS: pattern recognition, partial last squares regression, probability calculation, value system, adult development

Knowledge on the distribution of various levels of adult development in large populations is useful in many contexts, e.g. in reorganization processes and in the planning phase preceding comprehensive interventions in a large organizations, communities, regions or nations. Most existing methods for acquiring data on adult development are not suitable for such applications. Interviews require massive resources in terms of competent staff, time and money when applied on large populations. Another option is to use questionnaires that are relatively inexpensive both to produce and to distribute, but require particular qualities in order to generate representative results. Questionnaires with open-ended questions, such as sentence completion tests, necessitate considerable efforts by the respondents and are therefore likely to deliver poor response rates (Hansell, Ronchi, Sprarcino, & Stordtbeck, 1985; Truluck & Courtenay, 2002; Michiel P Westenberg, van Strien, & Drewes, 2001). Multiple choice-questions, on the other hand, are easy to complete for the respondent and are more likely to generate sufficient response frequencies. A limitation with questionnaires based on pre-formulated response alternatives is that they often cause noisy data that is difficult to deal with by conventional statistical methods.

One way of extracting meaningful information from noisy or complex data is to use multivariate statistics, e.g. principal component analysis. There are several different mathematical varieties of principal component analysis (for review, see Jolliffe 2002). The basic principle of all of them is to make data more interpretable through data reduction and to find latent structures in the data. Some methods are mathematically designed to recognize similarities in response patterns, i.e. pattern recognition methods. The later methods have become common in a wide spectrum of applications, e.g. recognition of speech, faces and finger prints, in diagnostics of tumors and other somatic diseases, and in social science, neuroscience and psychological research (e.g. Johansson et al. 1991; Duda et al 2001; Henningsson et al. 2002; Daerga et al. 2008; Brunelli 2009). In this paper a method called Partial Least Square Regression was used for pattern recognition analysis of value systems.

Table 1. The relative importance of the 38 variables, shown as variable loadings for the three significant components. The response alternatives of each item have been arranged in a sequential order. For a detailed description of the items and statements, see Appendix 1.

Items	Component 1	2	3
B5-responsibility for integration			
government and parliament	0,228	-0,101	0,191
national authorities	0,241	-0,063	0,246
local authorities	0,261	-0,143	0,155
non-profit organizations	0,255	-0,111	0,025
all Swedish citizens	0,277	-0,156	0,053
my own	0,281	-0,181	0,017
the refugees	0,154	0,059	0,168
C1-affinity			
my family	0,144	0,168	0,155
my friends	0,159	0,209	0,226
groups I share interests with	0,122	0,222	0,108
co-workers	0,110	0,262	0,120
local inhabitants	0,157	0,335	0,027
Swedes	0,114	0,314	-0,015
Europeans	0,182	0,195	0,053
all human beings	0,219	0,198	-0,073
refugees	0,282	-0,022	-0,087
C2-important			
satisfy my needs	-0,154	-0,050	0,237
social atmosphere	-0,034	-0,105	0,159
occupational skills	-0,112	0,151	0,134
develop others people	0,016	0,154	-0,340
develop myself	0,200	-0,037	-0,250
C3-good boss			
natural authority	-0,209	0,030	-0,014
social atmosphere	0,038	0,057	0,091
expert competency	-0,091	0,075	-0,036
focused on achievements	-0,096	0,167	-0,132
all co-workers are individuals	0,097	-0,082	0,008
social and global goals	0,118	-0,021	-0,133
C4-law and regulation			
difficult to satisfy my needs	-0,075	-0,050	-0,086
followed without exception	-0,064	0,327	0,020
to obtain a fair society	0,022	0,279	0,081
needs of the individual	0,100	-0,201	0,078
principally based	0,011	-0,224	-0,099
C5-thinking about			
satisfy my needs	-0,192	0,006	0,255
socially accepted	-0,117	-0,042	0,190
occupational skills	-0,042	0,000	0,215
efficiency and productivity	-0,032	0,140	-0,258
tolerant society	0,218	-0,064	-0,287
saving the mankind	0,147	0,072	-0,283

The aim of the present study was to use pattern recognition statistics to disclose characteristic response patterns in data generated by a multiple choice-questionnaire on psychological value systems, and to use probability calculations to categorize individual responses obtained from a random sample of adult. The characteristic response patterns were hierarchically arranged based on values that have been found to be common in different stages of ego development. The hierarchically arranged categories of value systems were validated by comparison with the Washington University Sentence Completion Test (ego development) and the Hierarchical Complexity Scoring System (complexity of reasoning).

» METHODS
The questionnaire

A questionnaire was constructed with the main aim to get information on views and attitudes on refugees, discrimination and the Swedish migration policy. This was a part of a larger evaluation of the refugee receptions in two Swedish municipalities.

To enable comparisons between people's views on refugees, discrimination and their general values systems, 6 out of a total of 21 questionnaire items were constructed to gain information on value system (Appendix 1). Our intention was to compose a number of issues and statements with the potential of discriminating between different conventional value systems. The response alternatives and the statements were selected to represent a sequence of values systems, ranging from pre-conventional or early conventional to late conventional or early post-conventional, that we believed should be differently appealing to various conventional levels of adult development. In the questionnaire these response alternatives and the statements were presented in a random order (Appendix 1).

The selection of items and statements were inspired by theories and empirical data presented by Cook-Greuter (1999), Loevinger and Blasi (1976), Loevinger & Hy (1996), Torbert (2004), and Westenberg et al. (1998), and selected and formulated to fit into the collective norm system of the Swedish society (e.g. distribution of responsibility, "glorification" of labor skills, work efficiency and expertise, and justice through equal opportunities). Three of the 6 questions dealt with rather general issues, i.e. affinity with different groups of people, issues regarded as personally important, and issues one often thought about, whereas three questions were linked to more specific topics, i.e. responsibility for integration of refugees, characteristics of a good boss and views on laws and regulations. For two of the questions the respondents were asked to grade, on an ordinal scale from 1 to 5, how much responsibility different authorities and groups have for integration of refugees, and how much affinity one feels for different groups of people. The instruction for the other four items was to rank 5-6 statements in accordance with how well or poorly the statements corresponded with the respondent's views and opinions. If none of the pre-formulated statements matched their most preferable view they were encouraged to formulate such a statement on the questionnaire (Appendix 1).

A total of 38 variables were derived from the 6 items and their concomitant response alternatives and statements (see Appendix 1).

The analyses

To identify prevalent response profiles among those who answer the questionnaire, i.e. to disclose distinct patterns of covariation over the 38 variables, pattern recognition statistics were applied. Partial Last Squares Regression (PLS), which is a type of principal component analysis, was selected since this method is relatively insensitive to multicollinearity in the data (Wold et al. 2001). In contrasts to conventional correlation statistics, PLS account for both linear and non-linear covariation between variables. In PLS the data table is represented as a swarm of points in a multidimentional space where the data points correspond to the individuals' response profiles in a coordinate system with as many axes as there are variables. A vector is fitted to the data swarm, by a least-square method that identifies the direction with the largest variation. Each data point is projected down on this line to create the individual scores of the first principal component (PC1). Based on the residual matrix a second vector, orthogonal to the first, is fitted to the data swarm, again to represent the direction with the largest variation. The projections on this second component (PC2) account for as much of the remaining variability as possible. This procedure can be repeated until the residuals are zero, but usually only the first few components are significant. The first PCs can be seen as uncorrelated variables that represent the most frequent response profiles in the data set.

The PLS also provides values of the so-called loading vectors, showing how the variables are combined to form the distribution of scores (individual response profiles). The relative weight of the individual variables on a PC is called variable loading. Some variables are more important than others. High positive and negative loading values indicate large importance on the distribution of scores (response profiles) whereas values close to zero indicate variables with marginal impact (see Table 1).

Before the PLS was commenced, the variables were scaled to unit variance by calculating the scaling weights as $1/SD_i$, where SD_i is the standard deviation of variable i over the objects. The statistical significance of the PCs was calculated by a cross validation method (Wold, 1978; Eastment & Krzanowski, 1982). In short, this technique implies that data of the matrix are pseudo-randomly selected and deleted, where after the incomplete matrix is used to calculate a PC. From this PC, the values from the deleted data are predicted. The matrix is then restored and new data randomly selected and deleted. This procedure is repeated until each data element has been deleted once and only once. The sum of the squared differences between the actual and the predicted values are taken as a measure of how adequately the PC predicts the data.

In the present study significant PCs reflect distinct response patterns common in the cross-section population, and it was assumed that these response profiles would correspond to different value systems. To estimate the correspondence between the individuals' response profile and the general response patterns disclosed by the PCs, probability scores were calculated. Thus, the three significant PCs of the first model generated three probability scores for each individual.

The probability score in relation to a given PC was calculated as follows. First the questionnaire score was multiplied, separately for each of the 38 variables ($scoreQ1 \ldots scoreQ38$), with the loading values of the corresponding variables ($loadV1 \ldots loadV38$). Thereby the questionnaire scores were scaled to the general response pattern disclosed by the PC. Then the sum of the scaled questionnaire scores was calculated ($sumPC1$).

$$\text{sumPC1} = \sum_{i=1}^{38} \text{scoreQ}_i \times \text{loadV}_i \qquad (1)$$

By relating the $sumPC1$ to the theoretical maximum and minimum of the scaled sum score for this particular component, a probability value ranging between 0 and 1 was obtained ($probPC1$).

$$\text{probPC1} = \frac{\text{sumPC1}}{\text{sumPC1}_{\text{max}} - \text{sumPC1}_{\text{min}}} \qquad (2)$$

This procedure was repeated for each of the significant components, thus showing the probabilities for an individual response profile to correspond with the significant profiles disclosed by the PLS. The classification of an individual's response profile, i.e. whether it showed the closes correspondence with the significant response profile identified by PC1, PC2 or PC3, was based on the largest probability value.

The population tested

The questionnaire was sent to a randomly selected adult population of 600 inhabitants of the municipalities of Vilhelmina and Åsele. These sparsely populated municipalities, located in the northwest of Sweden, hold populations characterized, in a national context, by a relatively low level of education, high unemployment rate and relatively large frequencies of blue collar workers.

After one reminder a total of 281 questionnaires were returned (response rate 47%). The gender and age distributions were rather even in the sample of returned questionnaires, i.e. 53% women (mean age, 54 years) and 46% men (mean age, 49 years).

Twenty-two of the questionnaires were too incompletely answered to be useful in the analyses. Another 32 did not satisfy the instructions on the ranking-questions, i.e. their responses were undifferentiated, or both incomplete and undifferentiated. For 27 of these at least 2 of the statements were ranked on 3 out of the 4 ranking-questions which eventually made them classifiable based on the general response profiles identified in the PLSs (see Figure 4). Thus, the PLSs were based on a sample of 227 acceptably completed questionnaires, whereas the sociodemographic characteristics of the different classes of value systems were compiled from 254 individuals.

The principal value systems identified through the PCs were sequentially ordered based on theories and empirical data presented by Cook-Greuter (1999), Loevinger & Blasi (1976), Loevinger & Hy (1996), Torbert (2004), and Westenberg et al. (1998).

Validation

To investigate to what extent the sequence of different value systems was related to other psychological domains of adult development, sub-groups of the study population were tested regarding their level of ego development and complexity of reasoning.

Ego development was assessed with the Washington University Sentence Completion Test (WUSCT). A Swedish translation of a 18-item version of the WUSCT was used (Rosén, 1997). The form was administrated to a sub-sample of the population (n=20). The response to each item was assigned a score between 1 and 10, and, for each individual, the total protocol rating was obtained through the automatic ogive rules (Loevinger & Hy, 1996). In accordance with the well established procedures, the scoring was made by one of the author (SK) who is an experienced scorer of WUSCT.

The value system questionnaire was constructed to permit inclusion of response alternatives and statements formulate by the respondents themselves. They were also encouraged to provide written feedback and comments on a separate page added to the questionnaire. Out of the 281 individuals who returned the questionnaire, 112 had included a new response alternative and/or statement and/or provided feedback and/or general comments. For a total of 47 individuals these written statements and comments were sufficiently extensive to permit analyses of the complexity of reasoning. The Hierarchical Complexity Scoring System (HCSS) was used to score the complexity of these statements and comments (Commons et al., 2007). The scoring, which was based on the highest level of reasoning shown by each individual, was done independently by two of the authors (PS and AE). Their indepen-

dent ratings were identical for 42 of the individuals (inter-rater reliability = 0.89). After discussions they reached consensus for three, but not for two of the individuals. For the latter two, the highest scores were used in the analyses.

Statistical tools

Calculations of mean values, correlation coefficients (Spearman´s rho, 2-tailed) and statistical significance of bivariate comparisons were performed with SPSS (version 11.5, SPSS Inc., USA). A multivariate program package was used for the PLS (SIMCA-P, version 11.0 Umetrics, Sweden).

» RESULTS
Partial last squares regression

The PLS of the 38 value-related items for the 227 individuals showed three statistically significant principal components, indicating three different response patterns. The model explained 33% of the variance in the data – 16.6% of the variance was explained by the first component, 9.1% by the second and 7.4% by the third.

The first component was defined by response profiles characterized by high scores on nearly all of the responsibility alternatives and by strong affinity with all people, including refugees, Europeans and all people (Table 1). Moreover, they ranked the following statements high: "it is important to meet other people in order to develop ones ability to critically scrutinize oneself and the norms of the society" (item C2), "a good boss should have long-term, social, humanistic and global visions (item C3), "laws should take into consideration the conditions and needs of the individuals" (item C4), and "how to achieve a more tolerant society" (item C5). These response profiles were also shaped by relatively low ranks of the statements focusing on self-satisfaction and social acceptance on items C2-C5.

Figure 1. Averaged response profiles for individuals classified with early, middle and late conventional value systems based on the PLS-model and probability calculations.
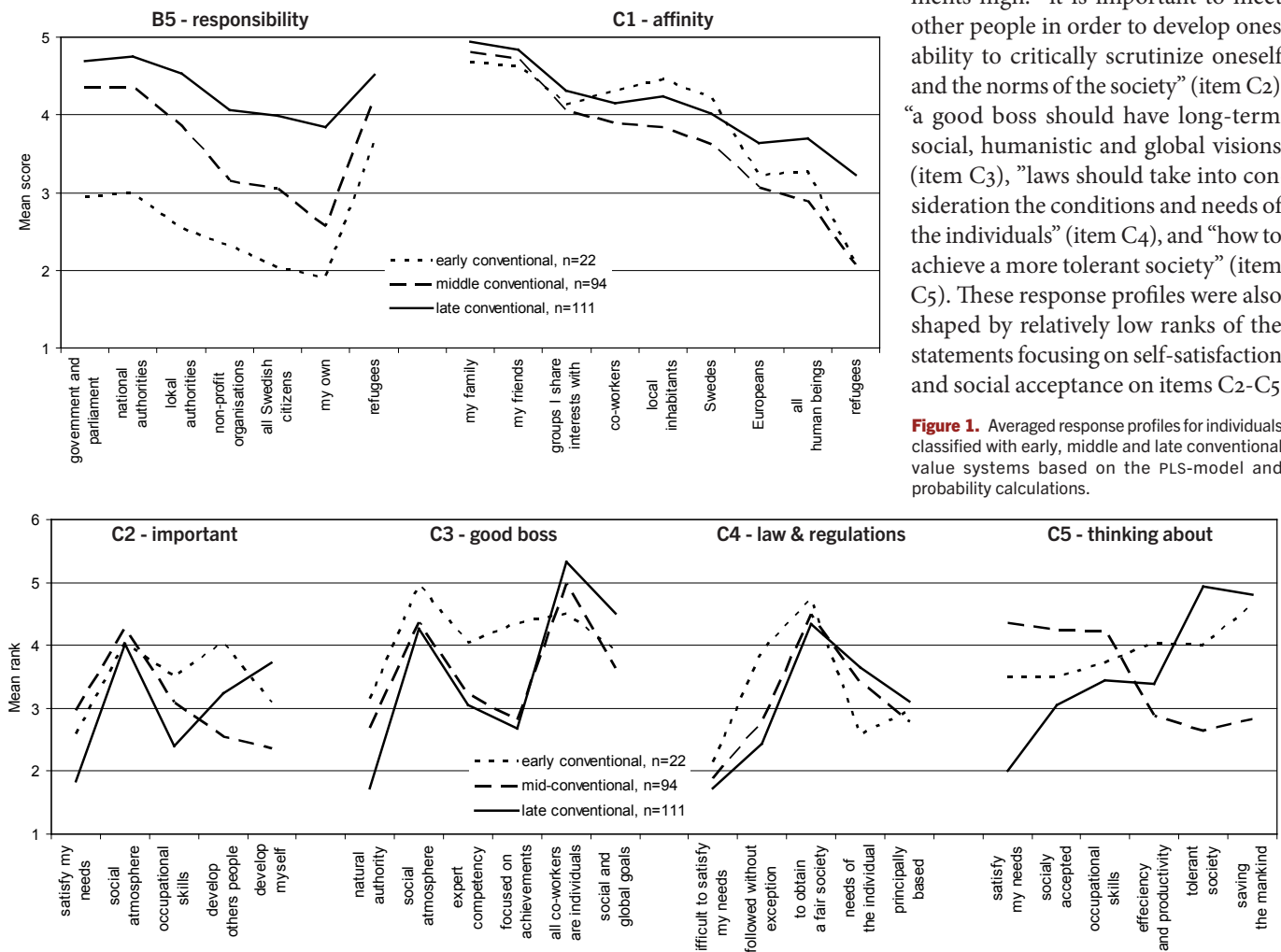
Table 2. Sociodemographic characteristics and responses on two questions on refugees for the sub-groups classified according to value system.

	Early conventional		Middle conventional				Late conventional					
			Sub-group 2		Sub-group 1		Sub-group 2		Sub-group 1		Sub-group 3	
	$n = 37$		$n = 25$		$n = 76$		$n = 12$		$n = 89$		$n = 15$	
Sociodemography	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men
Gender distribution	14%	16%	7%	12%	30%	30%	2%	7%	40%	30%	7%	5%
Mean age, years	68	49	33	42	44	44	75	39	52	56	61	60
Highest level of education												
» compulsory, 6-9 years	55%	17%	10%	27%	10%	22%	67%	22%	9%	26%	45%	50%
» college, 10-12 years	28%	71%	80%	40%	58%	51%	0%	67%	54%	46%	33%	33%
» university, >13 years	17%	12%	10%	33%	32%	27%	33%	11%	37%	29%	22%	17%
Occupation												
» working	17%	57%	55%	69%	62%	82%	33%	56%	61%	72%	46%	50%
» retired	72%	24%	9%	19%	19%	15%	67%	33%	25%	21%	45%	33%
» sick leave, unemployed	11%	10%	18%	6%	10%	0%	0%	0%	3%	5%	0%	0%
» studying	0%	9%	18%	6%	9%	3%	0%	11%	11%	2%	9%	17%
Questions on refugees												
*What is your view on the fact that Sweden accepts refugees?**												
Very positive	14%	8%	20%	7%	15%	9%	33%	50%	35%	15%	38%	50%
Very negative	43%	31%	10%	7%	0%	3%	0%	17%	0%	0%	0%	0%
Are refugees discriminated in Sweden?												
Yes	0%	17%	50%	53%	46%	50%	67%	67%	58%	56%	75%	60%
No	53%	61%	20%	13%	31%	17%	33%	11%	9%	29%	13%	20%
Don't know	47%	22%	30%	33%	23%	33%	0%	22%	33%	15%	13%	20%

* the response alternatives were given as a 5-graded ordinal scale, from 'very positive' (5) to 'very negative' (1)

The response profiles that shaped the second component were dominated by low scores on the responsibility alternatives, particularly "my own", and by strong affinity with "co-workers", "local inhabitants" and "Swedes", and weak affinity with "refugees" (Table 1). Among the ranking issues, this component was characterized by high ranking of the statements "value my expert competencies and occupational skills" and "developing other people to increase productivity and economic growth" (item C2), "a good boss should be focused on achievements and economic growth" (item C3), "laws and regulations should be followed without exceptions" and "laws and regulations are needed to obtain a fair society" (item C4), and "how to create more effective and productive companies and social structures" (item C5). This component was also formed by disagreement with "laws and regulations should consider the conditions and needs of the individuals" (item C4), and "laws and regulations could always be questioned, except those based on fundamental principles" (item C4).

The third significant component was dominated by response profiles claiming that the national authorities and the refugees should take large responsibilities for integration (Table 1). The affinity was strongest with family and friends, and weakest with "all human beings" and refugees. Among these profiles the statements focusing on self-satisfaction, social atmosphere and occupational skills, were highly ranked (items C2 and C5). Low rankings were given to the statements "contribute to developing other people in order to increase productivity and economic growth" and "to meet people with other values who develop my ability to critically scrutinize myself and the norms of society" (item C2). Other low ranked alternatives were "how to achieve a more tolerant society where all people and cultures are respected", "how mankind can be saved from global pollution, starvation and oppression" and "how to create more effective and productive companies and social structures" (item C5).

Classification of the test population

The classification procedure, i.e. the comparison of the probability values of the individual response profiles, demonstrated that 111 showed the best correspondence with the profile defined by the first principal component, 22 individual profiles with the second component, and 94 individual profiles with the third component.

As deemed from the average response profiles shown in Figure 1, the three groups seem to reflect differently developed value systems. The group with profiles corresponding to the first component appears to be more developed than the other two groups. These individuals indicate larger personal responsibility for integration, stronger affinity with people in general, lower "inclination/need" for self-satisfaction, deeper awareness of individuality, tolerance and personal development. This group seems to correspond to

Table 3. Comparison of classifications according to the ego development (WUSCT) and the value system.

Value system	Stage	N	WSCT	Stage	N
Early conventional	4	1	Conformist	4	3
Middle conventional					
Sub-groups 1 & 2	5	1	Self aware	5	4
Late conventional					
Sub-groups 1 & 2	6	14	Conscientious	6	10
Sub-groups 3	7	4	Individualist	7	3

late conventional or early post-conventional value systems, as described by others (Cook-Greuter, 1999; Loevinger & Blasi, 1976; Torbert, 2004), and is hereby denoted *late conventional*.

The other two groups demonstrate common characteristic in comparatively low personal responsibility for integration and poor feeling of affinity with people outside Sweden (particularly with "intruders" like refugees), and regard social acceptance and security as more important than individuality (Figure 1). But there are also some striking differences between there two groups. The group with profiles corresponding to the second component indicates very low responsibility scores for all authorities, Swedes and themselves, but not for the refugees who, by far, holds the largest

responsibility for their integration. They show strong beliefs in leaders who are "natural" authorities, and they hold the opinion that laws and norms should be followed without exceptions. They demonstrate the most nationalistic affinity pattern and eco the Swedish norm of glorifying expert skills, efficiency and productivity. The value system of this population seems to be largely in conformity with the collective norm system, as described by others (Cook-Greuter, 1999; Loevinger & Blasi, 1976; Torbert, 2004), and is hereby denoted *early conventional*.

The third group, which corresponds mostly with the third principal component, deviates significantly from the other two groups in some distinct ways. Their profile indicate the greatest need for self-satisfaction, the most frequent thoughts on how once occupational skills might be improved, in combination with the weakest understanding of the need for developing others, developing oneself, creating a tolerant society, and managing global social and environmental problems (Figure 1). This group seems to match the self-consciousness level of development, as described by others (Cook-Greuter, 1999; Loevinger & Blasi, 1976; Torbert, 2004), and is hereby denoted *middle conventional*.

To explore whether these three populations could be further separated into sub-groups with more specific characteristics, a new PLS was run on each of these. The analysis of the early conventional-group produced one significant principal component only. However, the middle conventional-group was split up into two separate groups as judged from the PLS and the probability calculation. The PLS-model explained a total of 21.4% of the variance of the data (12.4% and 9.0% explained variance for the two components). Most of the individual response profiles showed a higher probability to fit with the first component (n = 72), while a smaller sub-group matched the pattern exposed in the second component (n = 22).
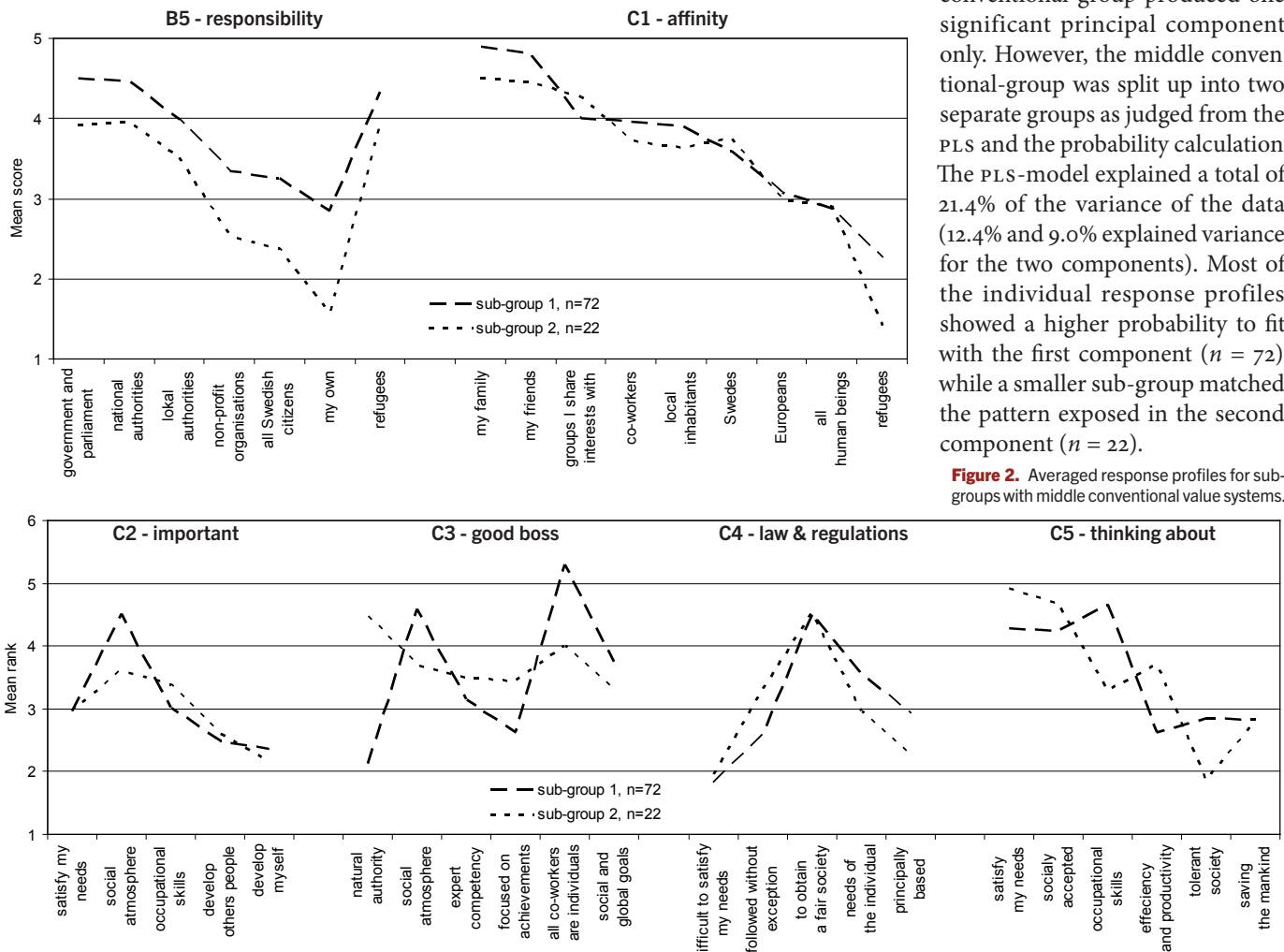
Figure 2. Averaged response profiles for sub-groups with middle conventional value systems.

The response profile of the larger mid-conventional sub-group confirms essentially with the characteristics described above (cf. Figure 1 and 2). However, the smaller sub-group showed some clearly deviating features. Overall this sub-group scored lower on the "occupational skill"-alternatives as compared with the other middle convetionals. The scores on the responsibility issues were generally lower, and particularly low for their own responsibility, the responsibility of non-profitable organizations and all Swedish citizens (Figure 2). Their affinity with refugees was weaker, and they regarded the social atmosphere as less important. The response profile on the good boss-issue was rather similar to that of the early conventional group except for higher scores on the natural authority-statement and somewhat lower rankings of the other alternatives. The profile on the law and regulation-issue was also more similar to that of the early conventional with the exception that scored lower on the statements "without exception" and "principally based". Their thoughts were most often directed towards self-satisfaction and to be socially accepted, while the statement "tolerant society" exhibited the lowest ranking. This sub-group could be in transition from an early to a mid-conventional value system.

The PLS on the late conventional group resulted in three significant principal components.

The model explained a total of 28.8% of the variance of the data (11.9%, 8.6% and 8.3% explained variance for the three

Table 4. Comparison of classification according to the complexity of reasoning (MCSS) and the value system.

Value system	n	%*	MHC concrete	abstract	formal	system	meta-system
Early conventional	5	14		1	2	2	
Middle conventional							
sub-group 1	15	20	1	2	9	3	
sub-group 2	3	12		1	2		
Late conventional							
sub-group 1	17	19		3	12		2
sub-group 2	1	8		1			
sub-group 3	5	33			2	2	1

Note. * the proportion classified in relation to the total size of each sub-group (see Figure 4).

components). The majority of the individual response profiles demonstrated the highest probability to fit with the first component ($n = 85$). Hence, the other two sub-groups were considerably smaller ($n = 11$ and $n = 15$, respectively).

Again, the response profile of the largest of the three sub-groups confirmed with the characteristics described above for the late conventional (cf. Figure 1 and 3). The sub-group identified through the second significant component showed response profiles somewhat
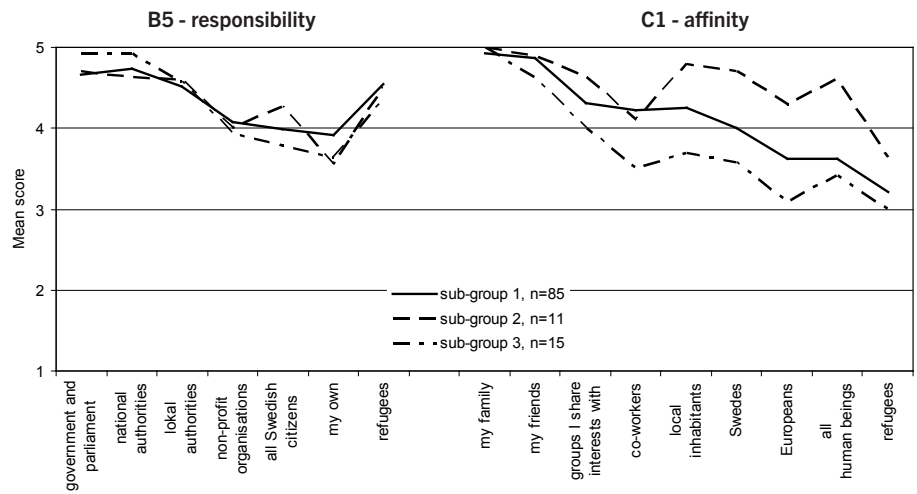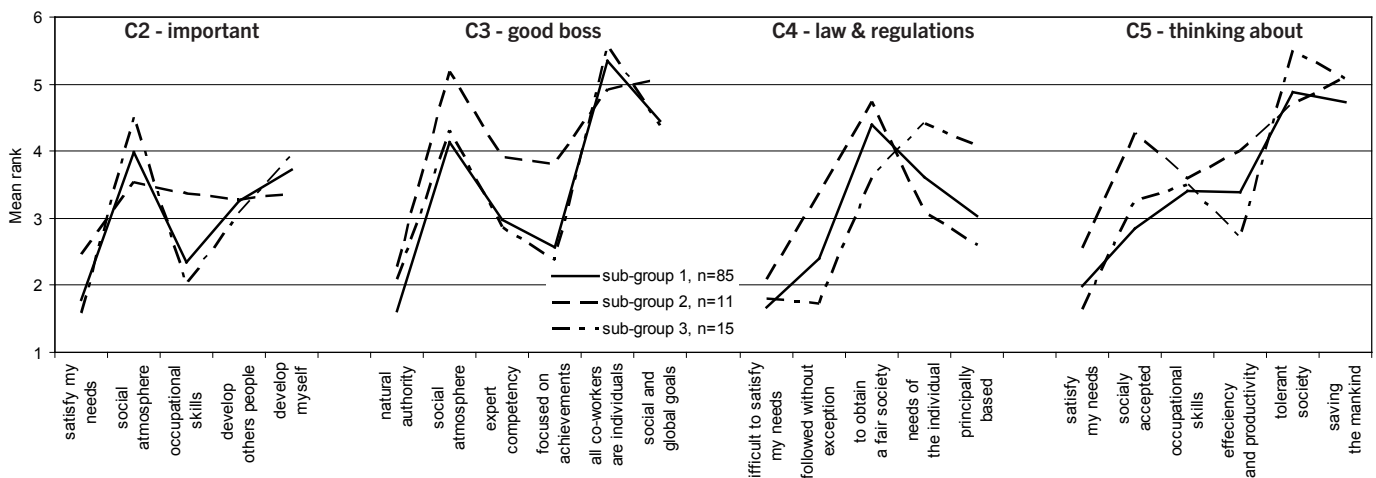

Figure 3. Averaged response profiles for sub-groups with late conventional value systems.
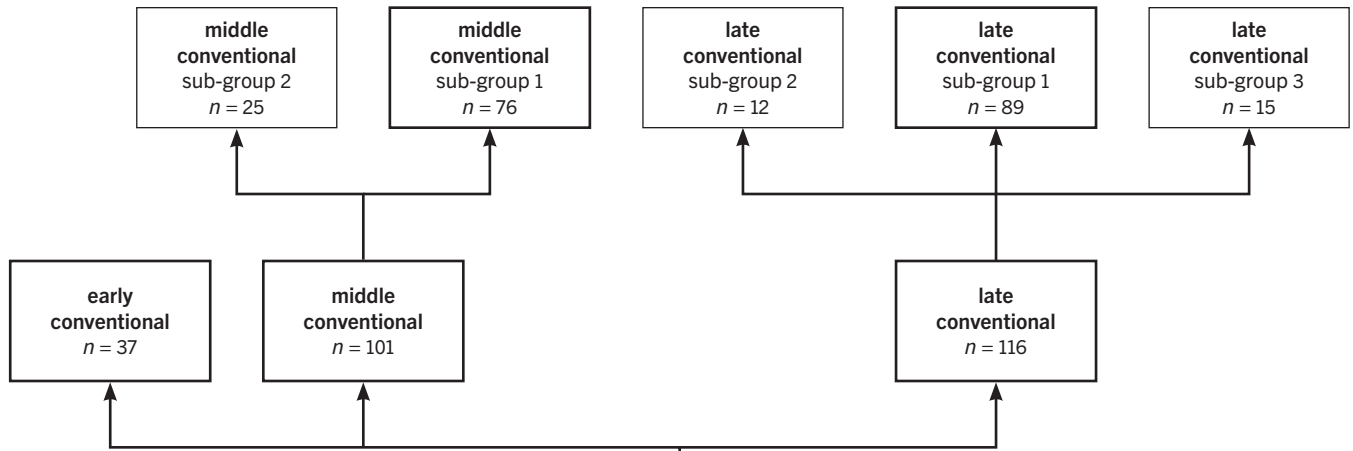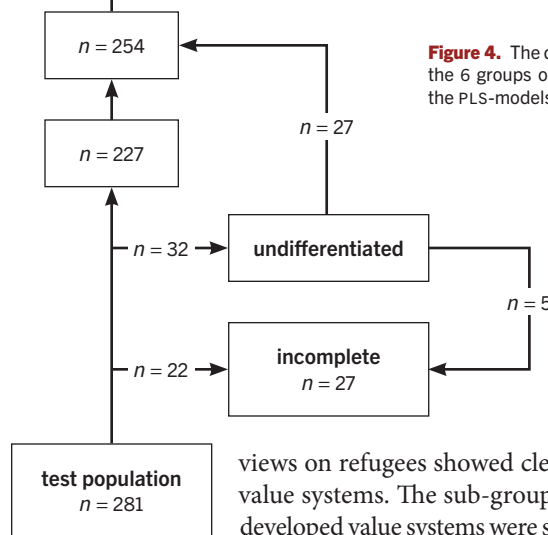
| middle conventional sub-group 2 n = 25 | middle conventional sub-group 1 n = 76 | late conventional sub-group 2 n = 12 | late conventional sub-group 1 n = 89 | late conventional sub-group 3 n = 15 |

| early conventional n = 37 | middle conventional n = 101 | | late conventional n = 116 |

n = 254
n = 227
n = 27

undifferentiated
n = 32
n = 5

incomplete n = 27
n = 22

test population n = 281

Figure 4. The distribution of individuals over the 6 groups of value systems disclosed by the PLS-models and probability calculations

similar to the middle conventional, except that they demonstrated higher responsibility and affinity scores (sub-group 2, Figure 3). They demonstrated higher scores in comparison with the main group of late conventional profiles on most of the statements concerning social atmosphere, occupational skills, efficiency and productivity. These profiles were also characterized by lower ranking of the individualistic alternatives. It seems reasonably to conclude that this sub-group contains individuals with value systems in-between the mid- and the late conventional value systems.

The third of the late conventional sub-groups showed response profiles with some distinct differences compared with the other sub-groups (sub-group 3, Figure 3). These individuals reported higher ratings on the individualistic alternatives on the "law and regulation" and on the "thinking about" issues. Their scores on the "saving the mankind" and "principally based laws" statements were also higher than the other sub-groups'. Moreover, they ranked "followed without exceptions", "to obtain a fair society" and "efficiency and productivity" alternatives lower than the other sub-groups. Their response profiles may indicate that this sub-group was composed of individuals in the border between late conventional and early post-conventional stages of value systems.

Based on the response profiles of the 6 identified groups it was possible to classify 27 of the 32 individuals who had provided too undiversified answers to be included in the PLS (see Methods). Fifteen of these were classified as having value systems typical for the early conventional group, 3 as mid-conventional sub-group 2 value systems, 4 as mid-conventional sub-group 1 value systems, 1 as late conventional sub-group 2 value system, and 4 as late conventional sub-group 1 value systems. Thus, a total of 254 individual were classified into a sequence of 6 different value systems (Figure 4).

Table 2 shows some sociodemographic data for the 6 sub-groups, together with their response characteristics on two of the refugee questions. There were no significant relations between the sequence of value systems, the gender distribution, the mean age, the level of education and the occupations, as deemed from non-significant bivariate correlations (p>0.05). However, the views on refugees showed clear-cut relations with the value systems. The sub-groups representing the more developed value systems were significantly more positive towards refugees in comparison with the sub-groups with less developed value systems (Table 2). The same trend was observed for the question on discrimination of refugees where the more developed value systems were associated with a higher frequency of recognition of discrimination whereas the less developed value systems showed a higher frequency of rejection of discrimination. For both these questions the responses were significantly correlated with the value systems (coefficients=0.361 and 0.278, respectively; p<0.001 for both).

Comparison of the value results with WUSCT and HCSS

A small group of the test population completed both the questionnaire and the 18-item WUSCT. To construct compatible classification scales, the 6 sub-groups identified by the questionnaire were ranked on the same ordinal scale as used for ego development (Table 3).

The classification scores obtained for value system and for ego development were significantly correlated ($r = 0.536$; $p = 0.015$). A perfect match between the two scoring systems was found for 11 out of the 20 individuals. In 6 cases the score on value system was one stage higher than that for ego development, while in two cases the score on value system was one stage lower. For one individual the score on value system was two stages higher than that for ego development.

The scores on ego development were not significantly correlated with age ($r = 0.331$; $p = 0.154$), level of education ($r = 0.151$; $p = 0.525$), attitude towards refugees (see Table 2; $r = 0.265$; $p = 0.259$), nor on the discrimination question (see Table 2; $r = 0.256$; $p = 0.289$).

A comparison of the stage of reasoning and the value systems for 47 individuals is shown in Table 4. The complexity ranged from concrete to meta-systematic, and the large majority of the individuals scored as formal (Table 4). However, the scores on the HCSS were not correlated with the value systems ($r = 0.034$; $p = 0.825$), neither with age ($r = 0.066$; $p = 0.665$) nor with the questions on attitude towards refugees and on discrimination (see Table 2; $r = 0.226$ and 0.169; $p = 0.136$ and 0.267 respectively), but significantly with level of education ($r = 0.353$; $p = 0.016$).

» DISCUSSION

This paper presents a method for disclosing value systems from questionnaires with pre-formulated response alternatives. The method is based on pattern recognition statistics and probability calculations that are applicable to any questionnaire containing a number of items where the response alternatives represent a sequential order e.g. of different levels of development within a psychological domain. The results from the population tested show that the method can disclose characteristic profiles of different value systems, and that these systems can be arranged in a hierarchical order similar to the levels of adult development identified through other methods. These methods might be a valuable tool to get information on the distribution of different levels of adult development in large populations, such as in communities and large organizations.

An important advantage with pattern recognition statistics and probability calculations is that this classification procedure is more objective than methods that depend on specifically trained raters or scorers. With the present approach there is no need to construct quantitative or qualitative classification criteria or 'ogive' rules that are bound to be more or less subjective due to e.g. semantic ambiguities, intra-disciplinary prejudices, cultural context, and the individual interpreter's stage of adult development. Partial least square regression identifies relations between all variables in the data, and by using the loading scores of each variable in the classification procedure all variables, but those with zero variance in all patterns, are taken into account. Most other methods applied in studies of adult development focus on one or a few qualitative features characteristic at different stages of development. By grouping the individuals according to the mathematic probability of belonging to identified response patterns the categorization procedure is unaffected by interpretation biases.

A most accurate method of assessing a person's level of psychological development is by talking to him/her, given that you know which questions to ask and that you are at least at the same level of development as the one you are speaking to. Thus, methods based on interviews hold a large potential to gain specific knowledge on various psychological characteristics. This has been evident by studies in a number of psychological domains showing convincing reliability and validity (e.g. King & Kitchener, 2004; Skoe & von der Lippe, 2002).

However, interviews are time consuming, expensive and require specific skills that make such techniques unsuitable for data acquisitioning in large surveys. Questionnaires with open-ended items require considerable efforts and cognitive abilities by the respondents. For instance, sentence completion tests are not likely to be completed by individuals who have linguistic dysfunctions. Distribution of sentence completion tests in large populations is destined to deliver poor response rates. Multiple choice-questions often generate larger response frequencies since they are easy to complete. Major drawbacks with this approach are that the outcome will be a direct reflection of the mind-frame of the constructor of the questions and response alternatives, and that the respondents' actual views may not be captured by the pre-formulated response alternatives. This limitation can be reduced by encourage the respondents to complement the response alternatives with own formulations, which was done in the questionnaire developed for the present study.

Another limitation with multiple choice-questionnaires is that they are inclined to generate noisy data with poor internal consistency. This is a considerable problem when the results obtained from single items are important for the study. The pattern recognition method applied in the present study identifies significant response profiles over a number of items, which makes the outcome less sensitive to noise and weak internal consistency. An implication of this approach is that conventional reliability and consistency tests, such as Cronbach's alpha and Rasch analyses, are not suitable to assess the reliability of the results. In PLS the reliability is evaluated through the statistical significance of the PCs and the part of the total variance that are explained by the model. In the PLS-models of the present study, the significant PCs together explained 21-33% of the total variance of the data. This is not too bad taken into account the large heterogeneity of the individual response profiles. Each individual showed a unique response profile, i.e. there were not two individuals with identical response profiles over the 38 variables. More importantly, more than one significant PC were defined in three out of the four PLS-models, indicating that the data set indeed contains several distinct response patterns with fundamentally different characteristics.

A prerequisite for identification and classification of response profiles through pattern recognition methods and probability calculations is that the questionnaire contains a sufficient number of questions with response alternatives that form a sequential structure. In the questionnaire used in this study the response alternatives and the statements were selected to represent a hierarchical sequence of values systems, ranging from pre-conventional/early conventional to late conventional/early post-conventional. In the analyses it became clear that some of the statements had poor discriminating power. Thus, the collection of response alternatives was certainly not the most optimal, but nonetheless indicates that the methods used for extraction and classification of response profiles is quite compelling. An exchange of some of the response alternatives in future versions of the questionnaire will probably reduce the noise and improve the discriminating potential.

In general, items where the respondent is forced to rank different statements are more discriminating than items where the response is marked on an ordinal scale (cf. questions C2-

C5 and B5, C1; Appendix 1). However, for some questions the ranking approach is unsuitable. An example is the affinity question (question C1) where nearly everyone would have top-ranked *'my family'* and *'my closest friends'*. These response alternatives cannot of course be omitted since that would have evoked serious doubts among the respondents about the credibility of the questionnaire as a whole. A restraint with the ranking approach is that the respondent requires a certain level of cognitive ability to understand and to complete the ranking task. In the population tested about 11% provided undifferentiated rankings, mostly by individuals with early conventional value systems (Figure 4), suggesting that the ranking task might be too challenging for individuals with poor capacity to deal with complex issues.

The first PLS disclosed three distinct response patterns that we believe relate to early, middle and late conventional value systems (Figure 1). The hierarchical structuring of these response patterns was based on their correspondence with characteristic features previously described for different conventional levels of ego development and action logic (Cook-Greuter 1999, Loevinger & Blasi 1976, Loevinger & Hy 1996, Torbert 2004, Westenberg et al. 1998). The early conventional value system bears resemblance to the diplomat/conformist stage, e.g. emphasizing social acceptance, natural authority, obedience to laws and regulations, echoing social norms such as the importance of occupational skills, efficiency and productivity, and disregarding individuality. The middle conventional value system shows similarities with the expert/self aware level, e.g. stresses self-satisfaction, individuality, social acceptance and occupational skills, but pays week attention to common and global issues, developing a tolerant society and a critical view on oneself and the society. The late conventional value system showed resemblance both to the achiever/conscientious stage and the individualist stage, e.g. pronounced self-responsibility, feeling of affinity with people in general, emphasizing a tolerant society, individuality, own development, global and social goals, relatively low need for self-satisfaction, social acceptance, and occupational skills.

The probability calculations allowed classification of the individual response profiles as predominantly early, middle or late conventional. Separate PLS-models based on these main sub-populations resulted in identification of three characteristic profiles within the late and two within the middle conventional sub-populations (Figure 2 and 3). However, the PLS on the early conventional sub-population produced only a single significant PC, hence indicating a rather homogenous response pattern. The altogether 6 identified response patterns were sequentially ordered, and the probabilities of the individual's response profiles to fit each of these were calculated. The distribution of the individuals over the 6 categories of value systems showed aggregations in the early conventional, in one of the middle and one of the late conventional value systems. In the sequential structure, the three other sub-groups of value systems were located in-between the larger ones, indicating transition stages between and/or sub-groups within the early, middle and late conventional and early post-conventional value systems (Figure 4).

The distribution of the test-population over the three main categories of value systems is in accordance what would be expected, and corresponds reasonably well with the distribution different stages of ego development reported in a mixed population from the USA (Torbert 2004). The relatively larger frequency of late conventional individuals in the present study is probably due to a sample bias. That is, people with late conventional value systems are probably the sub-group in which the response rate is the highest (there is a reason why this level of ego development is called conscientious). In contrast, pre-conventional people are the least likely to participate in any kind of voluntary surveys or study, unless they are offered a personal reward or are threatened to be punished. A contributing factor to the relatively larger group of individuals classified as late conventional is that this group might conceal post-conventional respondents. Although this is quite likely as indicated by the response pattern characteristic for the late conventional sub-group 3, this sub-group was very small (Figure 3). The questionnaire was not designed to identify value systems beyond the early post-conventional level since such individuals are very uncommon and even unique in the part of Sweden where the present survey was done (e.g. lack of companies and authorities that attract such people, local culture that strongly promotes conventional values, continuous emigration of 'unconventional' people etc).

The values hold by a person is a result of norms of the society and psychological properties. Societal norms are supposed to have a larger impact on the individuals' value system at conventional stages of development, particularly at the early stages, than in the post-conventional ones (Loevinger & Blasi, 1976, Cook-Greuter, 1999). In a psychological perspective the internalized values are influenced by, or the consequence of, a number of qualities such as e.g. the perception of oneself, others people and the world around us, identity and affinity, and cognitive abilities. The preliminary validation measures indicate that the value systems, defined by the selection of items compiled in the present questionnaire, are related to ego development but not to the complexity of reasoning.

The positive correlation observed between value priorities and WUSCT-scores was expected since the questionnaire was constructed to capture characteristic value priorities in different stages of ego development. Yet, the relatively weak correlation coefficient shows that the WUSCT and the value system test are different constructs. A significant association between stages of ego development and value systems has been reported previously (e.g. Helson & Wink, 1987).

The questions and response alternatives selected to investigate people's value priorities in the present study are to various extent related to three out of the four main domains of ego development; character development, interpersonal style and conscious preoccupations (Loevinger & Blasi, 1976; Loevinger, 1998). Attempts to include the fourth domain, cognitive style, were deliberately omitted since cognitive capacity is quite unfeasible to conclusively investigate by a questionnaire with pre-formulated response alternatives. This might partly explain the lack of correlation between the value profiles and the HCSS-scores.

Another potential reason for the absence of relation between the value systems and the HCSS-scores might be that the written statements and comments on with we scored complexity of reasoning poorly represented the individual's actual ability to handle complexity (i.e., a statistical type 2 error). The statements and comments covered a wide range of topics and some of the comments were clearly written in an emotional state of mind and perhaps therefore reflected an unrepresentatively low level of complexity. Also, the method of HCSS is vulnerable for misclassifications. Scoring the complexity of people's thinking or reasoning is unavoidably influenced by the scorer's presumptions, semantic interpretations and ability to identify different levels of complexity. For instance, a given word, abstraction or concept can be used by people at different stages of adult development, but their understanding and interpretation of it could differ significantly. Different scorer might interpret the logical structure of a given response rather differently, which is evident by the imperfect correlations obtained in inter-rater reliability studies of HCSS. To some extent the semantic ambiguity could be coped with by probing for people's actual understanding of a given word, abstraction, logical connection, coordination of systems etc. During interviews this could be done by asking probing questions. However, that is not an option when written statements from questionnaires are analyzed. An implication of this is that the complexity scoring presented in this paper sometimes contained a fair amount of 'reading behind the lines' and intuitive probability estimates. To minimize this bias more than half of the responses were omitted since we deemed them as inconclusive regarding their level of complexity.

A more likely explanation for the absence of a significant relation between value priorities and complexity scores is that cognitive abilities, like the complexity of reasoning, and value systems are developed independently, at least in conventional stages of development where value priorities largely constitute an echo of norms and values held by the in-group culture (Loevinger & Blasi, 1976). This suggestion is in agreement with previous studies reporting non-significant relationships between cognitive development and ego development (King et al., 1989; Commons et al., 1989). Within a given stage of complexity, using equally solid lines of arguments and logical connections, different persons might end up at diametrically different value priorities. For example, one might be against immigration since it implies that we have to use considerable societal resources before the immigrants have learned our language, got a job and can make a reasonable contribution to the society, while another might be in support of immigration since it provide the society with new citizens with competencies that are valuable at the labor market. Both are formal lines of augments but end up at completely different value priorities.

In late conventional and post-conventional stages of development, complexity and value scores are more likely to be positively related (Cook-Greuter 1999). This is supported by the observation that the three individuals who demonstrated meta-systematic reasoning in the present study showed late conventional or early post-conventional values profiles.

The poor correspondence between value profiles and the complexity of reasoning might raise doubts about whether value systems can be arranged in a hierarchical structure at all. However, we think that there are good reasons to believe that. The value profiles were positively correlated with the ego development scores and, more importantly, the hierarchically arranged sequence of value profiles conforms to general developmental principles (Sjölander 2013). At later stages of development the individuals were more consistent in their value priorities and in their capacity to differentiate between different values, their ability to coordinate and integrate different values increased, and their autonomy increased in relation to culturally important values. ■

REFERENCES

Brunelli, R. (2009). *Template matching techniques in computer vision: theory and practice*, New York: Wiley.

Commons, M. L., Armon, C., Richards, F. A., & Schrader, D. E. (1989). A multi-domain study of adult development. In M. L., Commons, J. D., Sinott, F. A., Richards, & C. Armon (Eds.), *Adult Development. Vol 1: Comparisons and applications of developmental models* (pp. 33-56). New York: Praeger.

Commons, M. L., Rodriguez, J. A., Miller, P. M., Ross, S. N., LoCicero, A., Goodheart, E. A., et al. (2007). *Applying the model of hierarchical complexity*. Unpublished manuscript.

Cook-Greuter, S. R. (1999). *Postautonomous ego development: a study of its nature and measurement.* Integral Publishers: Dissertation Series.

Daerga. L., Edin-Liljegren. A. & Sjölander. P. (2008). Quality of life in relations to physical, psychosocial and socio-economic conditions among reindeer-herding Sami. *International Journal of Circumpolar Health*, *67*(1),10-28. doi:10.3402/ijch.v67i1.18223

Dawson, T. L. (2002). A comparision of three developmental stage scoring systems. *Journal of Applied Measurement, 3*(2), 146-189.

Dawson, T. L. (2003). A stage is a stage is a stage: a direct comparison of two scoring systems. *Journal of Genetic Psychology, 164*(3), 335-364. doi: 10.1080/00221320309597987

Dawson, T. L. (2004). Assessing intellectual development: three approaches, one sequence. *Journal of Adult Development, 11*(2), 71-85. doi: 0.1023/B:JADE.0000024541.84265.04

Dawson, T. L. (2006). The meaning and measurement of conceptual development in adulthood. In C. H. Hoare (Ed.), *Handbook of adult development and learning* (pp. 433-454). Oxford: Oxford University Press.

Dawson, T. L., Xie, Y., & Wilson, M. (2003). Domain-general and domain-specific developmental assessments: do they measure the same thing? *Cognitive Development, 18*(1), 61-78. doi: 10.1016/S0885-2014(02)00162-4

Duda, R. O., Hart P. E., & Stork D. G. (2001). *Pattern classification* (2. ed). New York: Wiley.

Eastment, H., & Krzanowski, W. (1982). Cross validatory choice of the number of components from a principal component analysis. *Technometrics, 24*(1), 73-77. doi: 10.2307/1267581

Helson, R., & Wink, P. (1987). Two conceptions of maturity examined in the findings of a longitudinal study. *Journal of Personality and Social Psychology*, *53*(3), 531-541. doi: 10.1037/0022-3514.53.3.531

Hansell, S., Ronchi, D., Sparcino, J., & Stordtbeck, F. L. (1984). Ego Development Responses in Written Questionnires and Telephone Interviews. *Journal of Personality and Social Psychology, 47*(5), 1118-1128. doi: 10.1037/0022-3514.47.5.1118

Henningsson, M., Sundbom, E., Armelius, B. A., & Erdberg, P. (2001). PLS model building: a multivariate approach to personality test data. *Scandinavian Journal of Psychology*, *42*(5), 399-409. doi: 10.1111/1467-9450.00252

Hoare, C. H. (2006). *Handbook of adult development and learning.* Oxford: Oxford University Press.

Johansson, H., Sjölander, P., & Sojka, P. (1991). Fusimotor reflex profiles of individual triceps surae primary muscle spindle afferents assessed with multi-afferent recording technique. *Journal de Physiologie*, *85*(1), 6-19.

Jolliffe I. T. (2002). *Principal Component Analysis, Series: Springer series in statistics* (2. ed.) New York: Springer.

King, P. M., Kitchener, K. S., Wood, P. K., & Davison, M. L. (1989). Relationships across developmental lines: A longitudinal study of intellectual, moral and ego development. In M. L., Commons, J. D., Sinott, F. A., Richards, & C. Armon (Eds.), *Adult Development. Vol 1: Comparisons and applications of developmental models* (pp. 57-72). New York: Praeger.

King, P. M., & Kitchener, K. S. (2004). Reflective judgment: theory and research on the development of epistemic assumptions through adulthood, *Educational Psychologist, 39*(1), 5-18. doi: 10.1207/s15326985ep3901_2

Loevinger, J. (Ed.). (1998). *Technical foundations for measuring ego development*. Mahwah, New Jersey: Lawrence Erlbaum.

Loevinger, J., & Blasi, A. (1976). *Ego development: Conceptions and theories* (1. ed.). San Francisco: Jossey-Bass.

Loevinger, J., & Hy, L. X. (1996). *Measuring ego development* (2. ed.). Mahwah, New Jersey: Lawrence Erlbaum Erlbaum.

Rosén, A. S. (1997). Measuring ego development by sentence completion in the Swedish context. In A. M. Carlsson, A. Cederström & H. Janson (Eds.), *Research in Rorschach and projective methods* (pp. 85-92). Stockholm: Swedish Rorschach Society.

Sjölander, P. (2013). Hierarchical analyses of value systems based on evolutionary principles. Manuscript in preparation.

Skoe, E., & von der Lippe, A. L. (2002). Ego development and the ethics of care and justice: The relations among them revisited. *Journal of Personality, 70*(4), 485-508. doi: 10.1111/1467-6494.05016

Stein, Z., & Heikkinen, K. (2009). Models, metrics, and measurements in developmental psychology. *Integral Review, 5*(1), 4-24.

Torbert, W. R. (2004). *Action inquiry: the secret of timely and transforming leadership*. San Francisco, California: Berrett-Koehler.

Truluck, J. E., & Courtenay, B. C. (2002). Ego development and the influence of gender, age, and educational levels among older adults. *Educational Gerontology, 28*(4), 325-336. doi: 10.1080/036012702753590433

Westenberg, M. P., Blasi, A., & Cohn, L. D. (Eds.). (1998). *Personality development: theoretical, empirical, and clinical investigations of Loevinger's conceptions of ego development* Mahwah, New Jersey: Lawrence Erlbaum.

Westenberg, M. P., van Strien, S. D., & Drewes, M. J. (2001). Revised description and measurement of ego development in early adolescence: an artifact of the written procedure? *Journal of Early Adolescence, 21*(4), 470-493.

....

APPENDIX 1

The questionnaire items

B5. How much responsibility do you feel different people and organizations have for the integration of refugees?

	A great deal				None at all
	1	2	3	4	5
» The government and Parliament	☐	☐	☐	☐	☐
» All Swedish citizens	☐	☐	☐	☐	☐
» Local authorities (e.g. municipality, county administrative board and county council)	☐	☐	☐	☐	☐
» Non-profit organizations (e.g. sports associations, churches and the Red Cross)	☐	☐	☐	☐	☐
» The refugees themselves	☐	☐	☐	☐	☐
» National authorities (e.g. the Migration Board, the National Agency for Education and the Labor Market Board)	☐	☐	☐	☐	☐
» My own	☐	☐	☐	☐	☐

C1. What affinity do you feel with the following groups?

	Close				None
	1	2	3	4	5
» Europeans	☐	☐	☐	☐	☐
» My closest friends	☐	☐	☐	☐	☐
» The local inhabitants	☐	☐	☐	☐	☐
» All human beings	☐	☐	☐	☐	☐
» My family	☐	☐	☐	☐	☐
» Swedes	☐	☐	☐	☐	☐
» Groups with which I share interests or opinions	☐	☐	☐	☐	☐
» Refugees	☐	☐	☐	☐	☐
» Co-workers	☐	☐	☐	☐	☐

C2. What is most important to you?

Rank the following statements where 1 is most important, 2 is second most important etc.

_____ To contribute to developing other people in order to increase productivity and economic growth

_____ That the social atmosphere provides security by accepting and appreciating me

_____ To meet people with other values who develop my ability to critically scrutinize myself and the norms of society

_____ That my surroundings satisfy my needs

_____ That my surroundings value my expert competencies and occupational skills

_____ Own formulation: _____

C3. What do you think makes a good boss?

Rank the following statements where 1 corresponds best, 2 corresponds second best etc.
A good boss...

_____ Prioritizes mutual understanding and social atmosphere of togetherness

_____ Is focused on achievements and economic growth

_____ Is a natural authority who knows what is best for me

_____ Understands the importance of expert competency in order to create maximum quality

_____ Has long-term, social, humanistic and global goals and visions

_____ Understands that all co-workers are individuals with different opportunities and limitations

_____ Own formulation: _____

C4. How do you view laws and regulations?

Rank the following statements where 1 corresponds best, 2 corresponds second best etc.

_____ Laws and regulations could always be questioned, except those based on fundamental principles (e.g. all humans' equal worth)

_____ Laws and regulations are needed to obtain a fair society

_____ Laws and regulations must take into consideration the conditions and needs of the individual

_____ Laws and regulations make it more difficult for me to satisfy my needs

_____ Laws and regulations should be followed without exception

_____ Own formulation: _____

C5. When you are not thinking about practical everyday problems (taking care of the household and children etc.), what do you mostly think about?

Rank the following statements where 1 is what you think about the most, 2 what you think about second most etc.

_____ How mankind can be saved from global pollution, starvation and oppression

_____ How my occupational skills can be improved or be put to better use

_____ How to achieve a more tolerant society where all people and cultures are respected

_____ How I can satisfy my needs in the easiest way possible

_____ How to create more effective and productive companies and social structures

_____ How I can be accepted and find security in my social surroundings

_____ Own formulation: _____