Running head: CONCEPTUAL DEVELOPMENT IN ADULTHOOD

Conceptual Development in Adulthood

Theo Linda Dawson

Cognitive Science

Hampshire College

Amherst, MA 01002

(413) 222-0633

tdawson@hampshire.edu

Conceptual Development in Adulthood

Until recently, there was one basic approach to the study of conceptual development in adulthood. This approach, exemplified in the work of scholars like Kohlberg and his colleagues (Colby, Kohlberg, Gibbs, & Lieberman, 1983), Kitchener & King (1990), Armon (Armon & Dawson, 2002), and Perry (1970), involves collecting data from a more or less representative sample of individuals at more or less frequent intervals over the course of several years. Guided by cognitive developmental theory, their longitudinal results, and sometimes, philosophical categories, researchers construct stage definitions. Although these definitions are cast as descriptions of reasoning structures, they include descriptions of the conceptual content associated with each developmental level. Stage definitions, along with exemplars from the construction samples targeted in these studies, become the basis for scoring manuals, reifying the conceptual content associated with each stage.

Employing these methods, the study of conceptual development in a given domain effectively ends with the development of a scoring manual. In the Kohlbergian case, the conceptual content associated with moral stages is based on the reasoning of white males; in the Perry case it is based on the reasoning of Harvard undergraduates; in Armon's case it is based on the reasoning of a varied but small convenience sample; and in Kitchener and King's work it is based primarily on the reasoning of college students. The scoring manuals associated with these systems all involve, to a greater or lesser extent, matching performances with exemplars taken from the original studies. One consequence is that the stage sequences described by these scholars are vulnerable to accusations of cultural bias (Gilligan, 1977). Though their contributions to cognitive developmental theory have been enormous, the methods of these scholars result in an incomplete account of conceptual development, in that, given the small samples on which their scoring systems are based, they necessarily fail to distinguish adequately between *what* is known and *how* it is known.

Reacting to this problem, many scholars have abandoned the cognitive developmental approach altogether, moving toward narrative (Tappan, 1990) or ethnographic (Shweder, 1996) descriptions of domain knowledge. Unfortunately, in this work, either age or educational level generally stands in as a proxy for development. This may be defensible in research on childhood development, but the relationship between age and developmental level breaks down completely in late adolescence, and the relationship between educational attainment and cognitive development becomes considerably less deterministic after high school.

Domain-based developmental stage systems introduce several additional problems: (1) there are no agreed upon criteria for assessing how well these scoring systems function as measures of developmental stage; (2) their developers do not always agree about what kind of behaviors are evidence of the attainment of any given stage; and (3) it is not clear how the developmental levels in one system are related to the levels in the other systems, making it difficult to conduct cross-domain comparisons.

My response has been to design a methodological approach called *developmental maieutics*[i], a method for examining conceptual development that separates the measurement of developmental level from the assessment of conceptual content, thus making it possible to pose questions about their interrelations. This chapter surveys the literature on this method, examining its validity and describing how it has been employed in a range of knowledge domains. I begin with a brief discussion of the relationship between development and learning, which is followed by a discussion of the issue of measurement, which is of central importance in the work described here. I then describe the hierarchical complexity construct and examine the validity and reliability of the Hierarchical Complexity Scoring System. Finally, I discuss the relationship between complexity level and the meanings we construct, and discuss how what we have learned about this relationship is being applied in educational contexts.

Development and Learning

The focus of this volume is on the relation between adult development and learning. Heretofore my work has not explicitly addressed this relation, but the independent assessment of the structure of reasoning performances and their conceptual content that is inherent to the research methodology described here offers a perspective on this issue.

In the cognitive-developmental literature, *learning* often refers to processes involved in the acquisition of domain content or skills, whereas *development* refers to progressive changes in the form of thought or other behavior. Some researchers argue that learning and development can involve different processes (Feldman, 1995; Sinnott, 1994)[ii], whereas others argue that they are intertwined, and make the distinction primarily for heuristic or pragmatic reasons (Fischer & Granott, 1995; Smith, 1998). The work presented here falls into the second category. The independent analysis of hierarchical complexity and conceptual content that distinguishes developmental maieutics is conducted in order to highlight the complex interrelation of conceptual content (in part, what has been learned) and hierarchical complexity (structure). Research results generated with this method overwhelmingly confirm the interdependence of these constructs and their contributions to development. For this reason, when my colleagues and I employ the term *development*, we refer to both learning (knowledge acquisition) and structuring. Simply put, learning is viewed as the component of development that involves interaction with the external environment, and structuring refers to the way in which knowledge is organized by the individual.

Developmental progress is profoundly dependent upon knowledge acquisition (or learning). First, as will become clear, movement from one developmental level to the next (within a given knowledge domain) is highly unlikely until the attainment of a high level of conceptual differentiation at the developmental level at which an individual typically functions (Dawson & Wilson, in press). Conversely, individuals appear to be unable to learn certain concepts until appropriate structures are apparent in performance. For example, my colleagues and I found that students in a grade nine physical science program did not, following instruction, demonstrate any understanding of the concepts of kinetic and potential energy unless, prior to instruction, their constructions of the energy concept demonstrated a particular level of abstraction (Dawson, 2003c)[iii]. We are not the first researchers to make observations of this kind (Fischer & Bidell, 1998; Inhelder, Bovet, & Sinclair, 1967; Walker & Taylor, 1991).

Second, the development of an adequate range of conceptions at a given developmental level appears to be a prerequisite for ‘normal’ development to the subsequent level. For example, Fischer and his colleagues (Fischer & Ayoub, 1994; Fischer et al., 1997) have shown that a failure to develop appropriate concepts along one or more conceptual strands (such as conceptions of a *good self*), rather than stalling development, alter the outcome of development by producing distorted conceptions at higher levels. These distorted conceptions can interfere with optimal functioning. Findings like these highlight the interrelation of learning and structuring, making it clear that they are both essential components of cognitive development. One educational implication is that programs that focus primarily on promoting structural development without adequate attention to teaching content knowledge and skills may ultimately impoverish development by producing inadequate conceptions at the next level, while programs that focus primarily on teaching content knowledge and skills may slow developmental progress by failing to provide adequate opportunities for restructuring.

The complex interrelation of learning and structuring makes it impossible to draw a clear line between these two aspects of the developmental process. The conceptual content of a new developmental level is, at least in part, the result of restructuring the conceptual content of the previous level. On the other hand, new knowledge is often obtained through interactions with the external environment in a process that involves both learning and structuring. From a Piagetian (1985) perspective, new knowledge is either *assimilated* to existing structures or *accommodated* through restructuring. In both processes, some kind of structuring takes place. This means that the conceptual content of any performance is the product of both acquisition (learning) and structuring.

Measurement

It can be argued that most progress in science proceeds from hypothesis to measurement to hypothesis. Consequently, without the development, calibration, and improvement of measures, scientific progress is likely to stall. The story of how measurement permits rapid scientific advance can be illustrated through any number of examples. The effect of the measurement of temperature on our understanding of the chemical properties of lead provides one such example. The tale begins with an assortment of semi-mythical early scientists, who all initially agreed that lead only melts when it is very hot—much hotter than the temperature at which ice melts, and quite a bit cooler than the temperature at which iron melts. These observations, made repeatedly, resulted in the related hypotheses that *temperature* can be thought of as a unidimensional scale, and that lead melts at a particular temperature. However, before it was possible to test these hypotheses it was necessary to develop a standard for measuring temperature. This was not an easy task. Partly because early temperature measuring devices were poorly calibrated, and partly because different temperature measuring devices employed different scales, the temperature at which lead melted seemed to vary from device to device and context to context. Scientists divided into a number of ‘camps.’ One group, using a new thermometer, argued that lead melts at a lower temperature than originally thought by the group who developed the first thermometer. Another group argued that there are multiple pathways toward melting, which explained why the melting seemed to occur at different temperatures. The third and final group argued that a single variable, temperature, could not be abstracted from particular contexts, because lead appeared to melt at different temperatures in different contexts.

Only when a measure of temperature had been adequately developed and widely accepted did it become possible to observe that lead consistently melts at about 327º C. Armed with this knowledge, scientists asked what it is about lead that causes it to melt at this particular temperature. They then developed hypotheses about the factors contributing to this phenomenon, observing that changes in altitude or air pressure seemed to result in small differences in the melting temperature of lead. So, context did seem to play a role. In order to observe these differences more accurately, the measurement of temperature was further refined. The resulting observations provided information that ultimately contributed to an understanding of lead's molecular structure.

Although parts of this story are fictional, it is true that the concept of temperature developed over time, that its measurement has undergone considerable development (and continues to be developed), and that its measurement has greatly contributed to our understanding of the properties of lead. However, the value of the thermometer, as we all know, extends far beyond this particular use. The thermometer is a measure of temperature *in general*, meaning that it can be employed to measure temperature in an almost limitless range of substances and contexts. This generality is one of the hallmarks of measurement.

Good measurement requires (1) the identification of a unidimensional, content and context-independent trait, property, or quality (temperature, length, time); (2) a system for assessing amounts or levels of the trait, property, or quality in interval units; (3) determinations of the reliability and validity of the assessments; and finally (4) the calibration of a measure. A good thermometer has all of the qualities of a good measure. It is a well-calibrated instrument that can be employed to accurately and reliably measure a general, unidimensional trait across a wide range of contexts.

What if cognitive scientists had access to an accurate, valid, and reliable general measure of cognitive development, one that spanned the developmental continuum from birth through adulthood? What might be some of the implications for cognitive research and education?

First, such a metric could be employed to investigate conceptual development in any knowledge domain. Once reasoning performances were assigned to their place on the developmental dimension, they could be subjected to a variety of content analyses. Matrices of conceptual content by developmental level would likely reveal patterns of conceptual change that are very difficult to expose employing conventional methods. While individual growth can only be studied longitudinally (Singer & Willett, 2003), a developmental metric would make it possible to meaningfully examine inter-individual developmental trends in cross-sectional data, putting an end to the questionable practice of using age as a proxy for development (Dawson, Commons, & Wilson, manuscript submitted for publication).

Second, a developmental metric would also help to eliminate the effects of sample bias in accounts of conceptual development. Conventional accounts of conceptual development are generally constructed by examining the behavior of individuals in small longitudinal samples. These accounts then form the basis for developmental assessment systems that often confound developmental level and conceptual content, such that *particular* concepts come to be overly identified with a given developmental level (Dawson, in press-a; Dawson & Gabrielian, 2003; Dawson, Xie, & Wilson, 2003). Because a developmental metric would allow us to specify an individual's place on the developmental continuum without reference to the particular conceptual content of his or her reasoning performance, we would be able to examine the empirical relation between particular conceptual content and a given developmental level, making it possible to interpret that relation as part of an independent analysis.

Third, a developmental metric could be employed to describe conceptual development across the entire developmental continuum, producing seamless accounts of development that could be employed to inform our understanding of developmental processes as well as curriculum design, instruction, and assessment. Though they have contributed importantly to our understanding of conceptual development in science, current accounts of conceptual development in the sciences are generally piecemeal, either because the research targets a particular age-group, or the developmental model being employed—such as the novice/expert model—dictates the comparison of two extreme groups. Attempts to tie together isolated results are complicated by the lack of a strong and coherent developmental theory. An accurate and reliable developmental metric would lend much greater specificity to efforts of this kind.

Fourth, a domain-general measure of development would make it possible to meaningfully compare developmental progress across knowledge domains. It would be possible to create a developmental report card like that shown in Figure 1, in which Lucile's developmental progress in multiple subjects is traced over time. Rather than employing age-norms or comparisons of Lucile with other managers in her department, this report card charts her progress with respect to an external criterion and her own developmental history. Further elaborations to such a report card could include developmental zones, indicating the range of performance expected for promotion to a higher management level. As research data, report cards of this kind could provide valuable insights into developmental processes while simultaneously providing information about the strengths and weaknesses of employees, students, teachers, or programs. Employed widely, report cards of this kind could provide meaningful, consistent information about employees' and students' intellectual development for employers, parents, teachers, and college entrance committees.

________________insert Figure 1 about here________________

Fifth, an accurate developmental metric could be employed longitudinally to compare the developmental trajectories and pathways of individuals. Accounting for the developmental dimension would reduce the danger of conflating developmental and other effects. For example, much cross-cultural research employs age or grade as a proxy for development. When differences in conceptions are found, these are often attributed to cultural differences. This practice is highly questionable, because though age and developmental level are correlated, they are far from an identity, particularly in adulthood (Armon & Dawson, 1997; Dawson et al., manuscript submitted for publication; Fischer & Bidell, 1998).

Sixth, an accurate content-independent developmental metric could be employed to link results from existing research, providing a more coherent picture of conceptual development. Such a metric could act as a single developmental dimension along which existing research findings could be arranged and reassessed.

Seventh, an accurate and reliable content-independent developmental metric could theoretically be employed to inform curriculum development. The link between students' cognitive developmental level and the likelihood that they will profit from instruction is already well-established (Case & Okamoto, 1996; Cavallo, 1996; Germann, 1994; Lawson, Alkhoury, Benford, Clark, & Falconer, 2000; Lawson & Renner, 1975; Lawson & Thompson, 1988; Lawson & Weser, 1990; Renner & Marek, 1990; Shayer & Adey, 1993). By employing a developmental metric to study conceptual development, we would be able to specify the relation between cognitive development and conceptual learning in a particular content domain. This could provide support for a cognitive developmental pedagogy in the form of concrete knowledge about the way in which particular concepts develop (typically, idiosyncratically, and optimally).

Finally, as suggested in the 4th point, above, a developmental metric could be employed in assessment and evaluation. Research, assessment, evaluation, and curricula could be coordinated around a single developmental agenda as suggested by Wilson and Sloane (2000). The above discussion of Figure 1 illustrates one way in which a developmental metric could be employed in the workplace. Another possibility that suggests itself is in item development for standardized assessments. Detailed knowledge of the course of conceptual development within a given domain would theoretically permit the development of high quality distractors based on students' actual conceptions (Amir, Frankl, & Tamir, 1987). Still another possibility is the development of practical scoring rubrics for educators, in which any given level would represent the same developmental attainment, regardless of knowledge domain.

The Current State of Developmental Assessment

In developmental psychology, although we agree that development occurs, we have not yet agreed upon a general developmental 'trait.' In other words, we do not yet agree about what we are trying to measure (requirement 1). Needless to say, this means that we do not yet have a developmental measure. Or does it? In this paper, I argue that several researchers, during the last century, have identified the same latent, general, developmental dimension, and that it is now within our capabilities to develop a technology that can be employed to take measurements along that dimension. I further argue that this measure has the potential to transform lifespan developmental science, making it possible for us to produce accurate, reliable, and practical developmental assessments in any knowledge domain, to deepen our understanding of developmental processes, to develop rich descriptions of the many pathways of conceptual development, to understand the impact of context on development, to meaningfully compare developmental progress across domains of knowledge, and to design better learning environments.

But before I begin, I want to address one of the roadblocks to the acceptance of a general developmental measure. This is the persistent notion that complex psychological phenomena cannot adequately be described with any method that involves isolating and measuring abstract dimensions of performance. Much like the contextualists in the 'lead' allegory, proponents of this view reject the idea that any useful abstract, general dimensions can be identified and employed in developmental science. They view developmental processes as highly complex and fundamentally tied to particular knowledge domains or contexts. Fortunately, we now have many examples of abstract measures that are employed to aid in the description and prediction of even the most complex phenomena. One of the best modern examples is the description and prediction of weather patterns. Imagine the state of weather prediction if contextualism had prevailed in meteorology! In contradistinction to the contextualist position, I argue that it is impossible to adequately describe complex phenomena, including the role of context in their development, without identifying, isolating, and measuring abstract, general dimensions.

In the following sections I describe the developmental dimension of *hierarchical complexity*, address the validity and reliability of a measure of hierarchical complexity as a developmental ruler, and show how this ruler has recently been employed to describe adult conceptual development in the moral, evaluative, and epistemological domains. I then discuss methods my colleagues and I are currently developing to examine individual differences in conceptual development, and suggest connections between what we are learning about adult conceptual development and the design of educational interventions intended to promote development and learning.

Hierarchical Complexity

Most cognitive-developmental researchers agree that development across knowledge domains does not necessarily proceed at the same rate (Fischer & Bidell, 1998; Lourenco & Machado, 1996). However, there is still considerable disagreement about whether development across domains can be characterized in terms of a single, generalized process. Domain theorists argue that entirely different processes apply in different knowledge domains. Others, although they acknowledge that unique structures and processes are associated with particular domains, also argue that a single general developmental process applies across domains. Piaget called this process *reflective (or reflecting) abstraction*, through which the actions of one developmental level become the subject of the actions of the subsequent level. The product of reflective abstraction is *hierarchical integration*. In conceptual development, hierarchical integration is reflected in the concepts constructed at a new level by coordinating (or integrating) the conceptual elements of the prior level. These new concepts are said to be qualitatively different from the concepts of the previous level, in that they integrate earlier knowledge into a new form of knowledge. For example, notions of *play* and *learning* constructed at one level are integrated into a notion of *learning as play* at the next level. This new concept cannot be reduced to the original *play* and *learning* elements, because even their earlier meanings are transformed when they are integrated into the learning as play concept. This hierarchical integration is observable in performance in the form of *hierarchical complexity*.

The general developmental model employed here has been strongly influenced by Piaget's stage model, Fischer's (1980) skill theory, and Commons' General Stage Model (Commons, Richards, with Ruf, Armstrong-Roche, & Bretzius, 1984; Commons, Trudeau, Stein, Richards, & Krause, 1998). In fact, the orders of hierarchical complexity (complexity levels) described here are equivalent to Fischer's 13 skill levels and Commons first 13 stages. They are numbered from 0 to 12 and named as follows: (0) single reflexive actions, (1) reflexive mappings, (2) reflexive systems, (3) single sensorimotor schemes, (4) sensorimotor mappings, (5) sensorimotor systems, (6) single representations, (7) representational mappings, (8) representational systems, (9) single abstractions, (10) abstract mappings, (11) abstract systems, and (12) single principles. (Definitions are provided, below.)

Not only are there definitional correspondences between analogous developmental levels described by Commons (Commons, Richards, with Ruf et al., 1984; Commons et al., 1998) and Fischer (1980), there is empirical evidence of correspondences between levels assessed with the scoring system based on these sequences (the *Hierarchical Complexity Scoring System*) (Dawson, 2003b) and four domain-based systems. The latter include Kitchener and King's (Dawson, 2002b; Kitchener, Lynch, Fischer, & Wood, 1993) stages of reflective judgment, Armon's Good Life stages (Dawson, 2002a), Perry's epistemological positions, (Dawson, in press-a) and

Kohlberg's moral stages (Dawson, in press-b; Dawson & Gabrielian, 2003).

Scoring for hierarchical complexity. The scoring procedures employed with the Hierarchical Complexity Scoring System are partially derived from Commons's (Commons et al., 1995) and Rose & Fischer's (1989) assessment systems. This scoring system, like its predecessors, is designed to make it possible to assess the hierarchical complexity of a performance without reference to its *particular* conceptual content. Rather than making the claim that a person occupies a complexity level because he or she has, for example, elaborated a particular conception of justice, the Hierarchical Complexity Scoring System permits us to identify performances at a particular complexity level and then to ask (empirically) what the range of justice conceptions are at that complexity level. Thus, it avoids much of the circularity[iv] of many stage scoring systems, which define stages in terms of particular conceptual content and domain-specific structures like social perspective-taking (Brainerd, 1993).

It is possible to score the hierarchical complexity of text performances because hierarchical complexity is reflected in two aspects of performance that can be abstracted from particular conceptual content. These are (a) *hierarchical order of abstraction* and (b) the logical organization of arguments. Hierarchical order of abstraction is observable in texts because new concepts are formed at each complexity level as the operations of the previous complexity level are "summarized" into single constructs. Halford (1999) suggested that this summarizing or "chunking" makes advanced forms of thought possible by reducing the number of elements that must be simultaneously coordinated, freeing up processing space and making it possible to produce an argument or conceptualization at a higher complexity level. Interestingly, at the single reflexive actions, single sensorimotor schemes, single representations, single abstractions, and single principles complexity levels, the new concepts not only coordinate or modify constructions from the previous complexity level, they are qualitatively distinct conceptual forms—reflexes, schemes, representations, abstractions, and principles, respectively (Fischer, 1980; Fischer & Bidell, 1998). The appearance of each of these conceptual forms ushers in three repeating logical forms—single elements, mappings or relations, and systems. Because these three logical forms are repeated several times throughout the course of development, it is only by pairing a logical form with a hierarchical order of abstraction that a rater can make an accurate assessment of the complexity level of a performance. For example, the statement, "In a good education, you get to have recess so you can play with your friends," is structurally identical to the statement, "In a good education, you get to socialize so you can learn how to relate to other people." Both are mappings. The first sentence, because its conceptual elements are representations, is a representational mapping. The second sentence, because its conceptual elements are abstractions, is an abstract mapping. Without the distinction between representations and abstractions it would be difficult to accurately score these texts. Other researchers have observed and described similar conceptual forms and repeating logical structures (Case, Okamoto, Henderson, & McKeough, 1993; Fischer & Bidell, 1998; Overton, Ward, Noveck, & Black, 1987; Piaget & Garcia, 1989).

Table 1 presents all 13 complexity levels, along with short descriptions of their logical structure and hierarchical order of abstraction (conceptual structure). Note that logical and conceptual structures are definitionally identical. We make a distinction between the two types of structure for heuristic and pragmatic reasons. When scoring texts, hierarchical order of abstraction refers primarily to the structure of the elements of arguments, which often must be *inferred* from their meaning in context, whereas logical structure refers to the *explicit* way in which these elements are coordinated in a given text. Examples of how we work with these two constructs are offered in the following section.

________________insert Table 1 about here________________

Complexity levels defined. Only the four complexity levels (single abstractions to single principles) commonly identified in adult performances are included in the following definitions. All of the examples provided in these definitions are from Dawson and Gabrielian's (2003) analysis of the conceptions of authority and contract associated with complexity levels in a sample of 747 moral judgment interviews scored with the Hierarchical Complexity Scoring System.

At the single abstractions level (abbreviated as "SA" in tables and figures), the new concepts are referred to as first order abstractions. These coordinate third order representations, which are equivalent to representational systems (the constructions of the previous complexity level). For example, the concept of *trustworthiness*, articulated for the first time at the single abstractions level, defines those qualities that make a person trustworthy rather than describing situations in which trust is felt or not felt. It is composed of qualities that produce trust, such as telling the truth, keeping secrets, and keeping promises. "It's always nice… to be trustworthy. Because then, if [someone has] a secret, they can come and talk to you." Concepts like *kindness, keeping your word, respect,* and *guilt* are also rare before the single abstractions level. "If you don't do something you promise, you'll feel really guilty." The most complex logical structure of this complexity level often identifies one aspect of a single abstraction, as in “Making a promise is giving your word” in which giving one’s word is an “aspect” of a promise.

Figure 2 portrays a visual representation, in the form of a concept map, of a 54-year-old[v] respondent’s argument about why promises should be kept. The respondent argues that a person should keep a promise because keeping promises is “the right thing to do.” When probed, the respondent comes up with three separate (uncoordinated) reasons for keeping promises, because people expect promises to be kept, because “people will trust you” if you keep a promise, and because “you might feel guilty if you break a promise.” All three of these reasons for keeping promises are considered to be first order abstractions, because they extract general abstract notions by coordinating concepts that appear for the first time at the representational systems level (Dawson & Gabrielian, 2003). Keeping promises will create trust, *in general*; people, *in general*, have expectations when promises are made; and breaking promises can produce negative emotional consequences, *in general,* for the promise-breaker. It is important to keep in mind that the *particular* concepts expressed by a respondent are important only to the extent that they embody a particular hierarchical order of abstraction. A rater must ‘look through’ the meaning of a particular conceptual element to abstract its hierarchical order of abstraction.

________________insert Figure 2 about here________________

At the abstract mappings level (abbreviated as “AM” in tables and figures), the new concepts are referred to as second order abstractions. These coordinate or modify abstractions. For example, the second order abstraction *basis* can be employed to coordinate the elements essential to a good relationship. “To me, [trust and respect are] the basis of a relationship, and without them you really don't have one.” Because they are usually employed to coordinate abstractions, concepts like *coming to an agreement, making a commitment, building trust,* and *compromise* also are rare before the abstract mappings level. “I think [Joe and his father] could come to an agreement or compromise that they are both comfortable with.” The most complex logical structure of this complexity level coordinates one aspect of two or more abstractions, as in “Joe has a right to go to camp because his father said he could go if he saved up the money, and Joe lived up to his commitment.” Here, Joe's fulfillment of his father's conditions determines whether Joe has a right or does not have a right to go to camp.

Figure 3 provides a map of the performance of a 58-year-old male, who provides three reasons for keeping promises. There are two mappings in this performance. The first is the assertion that “broken promises can harm relationships because they cause pain and reduce trust.” This mapping coordinates two abstract consequences of promise-breaking into the general notion that broken promises do harm to relationships. The second is the assertion that keeping promises makes it possible for people to “depend on one another.” This mapping coordinates the perspectives of at least two individuals to form the notion that keeping promises produces mutual benefits. Note how this idea builds on the single abstractions notion that people will trust you if you keep promises.

________________insert Figure 3 about here________________

At the abstract systems level (abbreviated as “AS” in tables and figures), the new concepts are referred to as third order abstractions. These coordinate elements of abstract systems. For example, the notion of *personal integrity* is rare before the abstract systems level, because it usually coordinates multiple abstract conceptions such as fairness, trustworthiness, honesty, preservation of the golden rule, etc., which are understood as interrelated aspects of the self. "[You should keep your word] for your own integrity. For your own self-worth, really. Just to always be the kind of person that you would want to be dealing with." Because they usually coordinate elements of abstract systems, concepts like *verbal contract, moral commitment, functional, development, social structure,* and *foundation* are also uncommon before the abstract systems level. "A promise is the verbal contract, the moral commitment that the father made to his son. It is the only way for the child to … develop his moral thinking — from watching his parent's moral attitude." The most complex logical structure of this complexity level coordinates multiple aspects of two or more abstractions. "Following through with his commitment and actually experiencing camp combine to promote Joe's growth and development, not just physically but psychologically, emotionally, and spiritually." Here multiple facets of Joe's personal development are promoted when he both keeps his commitment and accomplishes his goal.

Figure 4 provides a map of the performance of a 51-year-old female. The respondent describes a system in which promise-keeping is both obligatory and sometimes impossible, "due to unforeseen circumstances." The reason for keeping promises is that one must stand by one's commitments. Doing so not only preserves one's personal integrity, but also builds a sense of trust, "which keeps society functioning." The notion of standing by one's commitments, the idea that doing so preserves one's integrity, the argument that the sense of trust built through promise-keeping keeps society functioning, and the notion of unforeseen circumstances are all examples of second order abstractions. Note how the notion that the trust built from promise keeping keeps society functioning (even in the presence of the effects of unforeseen circumstances) builds on the abstract mappings idea that keeping promises makes it possible for people to depend on one another.

________________insert Figure 4 about here________________

At the single principles level (abbreviated as "SP" in tables and figures), the new concepts are referred to as first order principles. These coordinate abstract systems. An elaborated notion of the *social contract*, for example, results from the coordination of human interests (in which individual human beings are treated as systems). "Everybody wants to be treated equally and have a sense of fair play. Because this is so, we have an obligation to one another to enter into a social contract that optimizes equality and fairness." Because, they are usually employed to coordinate abstract systems or emerge from the coordination of abstract systems, concepts like *autonomy, fair play, heteronomy, higher order principle,* and *philosophical principle* are rare before the single principles level. "The only time we're justified in breaking the social contract is when a higher principle, such as the right to life, intervenes." The most complex logical structure of this complexity level often identifies one aspect of a principle or axiom coordinating systems, as in "Contracts are articulations of a unique human quality, mutual trust, which coordinates human relations." Here, contracts are seen as the instantiation of a broader principle coordinating human interactions.

Figure 5 presents a map of the performance of a 57-year-old male. Here, "mutual trust" is employed as a single principle supporting an argument for keeping promises. The rationale for employing this principle is that "most social conventions" and "all moral principles" are based on trust. Both "all moral principles" and "most social conventions" are third order abstractions. Note how this single principles argument builds on the abstract systems notion that trust keeps society functioning.

________________insert Figure 5 about here________________

Validation of the Hierarchical Complexity Scoring System: Hierarchical complexity scoring can be conducted in any knowledge domain because hierarchical order of abstraction and logical structure guide scoring, rather than the identification of *particular* conceptual content as in conventional domain-based systems. The domain generality of hierarchical complexity scoring raises the question of whether the Hierarchical Complexity Scoring System assesses the same dimension of performance as conventional developmental assessment systems do.

Five validation studies have shown that the Hierarchical Complexity Scoring System and one of its predecessors, the General Stage Scoring System (Commons et al., 1995) predominantly assess the same dimension of performance as more content-dependent stage-scoring systems. In the first of these, a think-aloud procedure was employed to compare the scoring behavior of five raters trained in the General Stage Scoring System with the scoring behavior of three raters trained in Kohlberg's Standard Issue Scoring System (Dawson, 2001). All raters scored the same 43 texts. A mean score for each text was calculated for each group of raters, resulting in two scores for each text, one based on the ratings of General Stage Scoring System raters and one based on the ratings of Standard Issue Scoring System raters. Despite the fact that the raters trained in the Standard Issue Scoring System justified their stage assignments on the basis of particular moral conceptions and interpersonal perspectives, whereas General Stage Scoring System raters justified their complexity level assignments in terms of logical structure, these mean scores were within one complexity level of one another 95% of the time (r = .94).

In a second study, the Hierarchical Complexity Scoring System, Armon's (1984) Good Life Scoring System, and the Standard Issue Scoring System were employed to score three different interviews administered to 209 5- to 86- year-olds. Correlations of .90 and .92 were found between the results obtained with the Hierarchical Complexity Scoring System and the Standard Issue and Good Life Scoring Systems (Dawson, 2002a). Dawson argued that these correlations, combined with patterns in the acquisition of analogous good life stages, moral stages, and complexity levels, provide evidence that the three scoring systems predominantly assess the same latent dimension: hierarchical complexity.

In a third study, Dawson, Xie and Wilson (2003) conducted a multidimensional partial credit Rasch analysis of the relationship between scores obtained with the Standard Issue Scoring System and scores obtained with the Hierarchical Complexity Scoring System on 378 moral judgment interviews from respondents aged 5 to 86. They found a correlation of .92 between scores awarded with the two scoring systems. This strong correlation suggests that these two systems primarily assess the same dimension of performance. However, the Hierarchical Complexity Scoring System revealed more stage-like patterns of performance than the Standard Issue Scoring System, including evidence of developmental spurts and plateaus.

In the fourth study, the relationship between complexity levels and Kitchener and King's (1990) Reflective Judgment Stages was examined (Dawson, 2002b). In a sample of 209 interviews of adolescents and adults, the correlation between complexity level scores and Reflective Judgment scores was .84. Agreement between Reflective Judgment scores and complexity level scores was within one Reflective Judgment stage 90% of the time. This is higher than the reported median inter-rater agreement rate of 77% within one Reflective Judgment stage (Kitchener & King, 1990).

In a fifth study, Dawson (in press-a) examined the relationship between scores obtained with the Hierarchical Complexity Scoring System and a Perry-based scoring system similar to that described by Mentkowski, Moeser, and Strait (1983). Scores agreed within one Perry level 82% to 98% of the time, depending on the levels that were compared.

Two additional studies examined the validity of the Hierarchical Complexity Scoring System as a measure of cognitive development. In these lifespan studies Rasch scaling was employed to examine patterns of performance in a set of 747 moral judgment interviews (Dawson, Commons, & Wilson, in review) and a set of 246 interviews addressing the question, “What is a good education (Dawson, manuscript submitted for publication)?” The age range of both samples was from 5 to 86. The authors examined these data for evidence supporting the specified developmental sequence as well as evidence for qualitative (as opposed to cumulative) change. They (a) identified 6 developmental stages in this age-range; (b) showed that performances were all either consolidated at a single complexity level or combined two adjacent complexity levels, a pattern that supports the specified order of acquisition; (c) demonstrated that movement from complexity level to complexity level proceeded in a remarkably consistent series of spurts and plateaus across the six complexity levels, reflecting the tendency for individuals to spend less time in transition from one complexity level to another than in periods of consolidation; and (d) showed that the task demands of moving from one complexity level to the next are remarkably similar, regardless of one's position in the developmental hierarchy. This means that the distance from one complexity level to the next is the same (in terms of task demands); consequently, the hierarchical complexity scale can legitimately be treated as an interval scale. This satisfies the second requirement for good measurement, a system for assessing the amount of a given trait in interval units.

In both studies the authors also showed that two of the complexity levels—abstract systems and single principles—were unlikely to occur before adulthood, and that patterns of performance on these two complexity levels were virtually identical to patterns of performance on the complexity levels found primarily in childhood and adolescence. This finding, which indicates that the developmental process is similar in childhood and adulthood, supports the claim that the Hierarchical Complexity Scoring System assesses a unidimensional developmental trait. This satisfies the first requirement for good measurement, the identification of unidimensional, context-independent trait.

It is notable that domain-based scoring systems have not revealed developmental patterns that are as unambiguously consistent with the requirements of measurement or with the postulates of stage-developmental theory as the results reported here.

Complexity Level and Meaning

Although the evidence presented thus far demonstrates convergent and construct validity for the Hierarchical Complexity Scoring System, it does not directly demonstrate how a measure of hierarchical complexity can help researchers to understand the development of meaning. The hierarchical complexity of a performance tells us a great deal about its form and hierarchical order of abstraction, but it makes no direct reference to its *specific* conceptual content. To develop an understanding of the progress of conceptual development in a given domain, specific conceptual content must be assessed independently, then reintegrated with hierarchical complexity information. Because complexity levels are assessed independently of *particular* conceptual content, it is possible to address questions about the relation between complexity level and meaning. This is impossible with conventional scoring systems because developmental level and particular meanings are conflated in these systems. The independent assessment of complexity level and conceptual content also makes it possible to address questions about individual (or cultural) differences in the behavior of individuals performing at the same level.

It is precisely because scoring with the Hierarchical Complexity Scoring System does not rely on the identification of particular conceptual content that it is possible to conduct separate analyses of complexity level and meaning. My colleagues and I have conducted (and continue to conduct) exhaustive analyses of the conceptual content of interview data in a number of knowledge domains, including moral reasoning (Dawson & Gabrielian, 2003), epistemological reasoning (Dawson, in press-a), evaluative reasoning (Dawson, manuscript submitted for publication), and, more recently, leadership reasoning, critical thinking, and science reasoning. In all of this work, we are guided by a methodological approach called *developmental maieutics*, which has been designed to expose patterns in conceptual development by submitting the same data to multiple forms of analysis (structural, conceptual, lexical) and then integrating the results of these analyses. In the abovementioned work, complexity coding of samples of texts was accompanied by independent and exhaustive content analyses, which we call *propositional* analyses. The first step in conducting a propositional analysis is to identify and code every relevant proposition made by each respondent represented in a given sample. So, for example, in a set of moral judgment interviews dealing with the concept of *contract,* we code every proposition about the nature of contracts. This coding is conducted with great attention to subtle differences in meaning. The results are then arranged in a complexity level by concept code matrix, which reveals patterns in the emergence of meaning within the knowledge domain being investigated. Analysis of these patterns allows us (1) to describe the conceptual content associated with each complexity level, and (2) to examine the ways in which the conceptual knowledge of one complexity level is reintegrated into new conceptions at the following complexity level. In other reports, Dawson and her colleagues (1998; manuscript submitted for publication; Dawson & Gabrielian, 2003) have described the process of developing accounts of conceptual development in greater detail.

Table 2 (adapted from Dawson, 1998) provides an example of a complexity level by concept code matrix (after considerable collapsing of the original concept code categories) related to the theme, *education as preparation for the good life*. The contents of tables like this one, combined with frequent references to the original texts, provide the basis for constructing rich descriptions of conceptual development. Tables like this one can also be employed to examine thematic trends across complexity levels. For example, 57.1% of individuals performing at the single abstractions level and 44.7% of individuals performing at the abstract mappings level, argue that a good education should prepare one for work or a job, whereas only 18.5% of individuals performing at the abstract systems level and 13.2% of individuals performing at the single principles level, make a similar argument. This suggests that the salience of preparation for work is diminished at the higher complexity levels. In contrast, a large percentage of individuals performing at the single principles level emphasize the development of social skills or ethical qualities relative to those performing at lower complexity levels, suggesting that the development of persons gains in salience at the highest complexity level.

________________insert Table 2 about here________________

Table 3 provides a summary of the conceptual content of interviews about learning, truth, and contracts/promises as they unfold from the single abstractions to single principles complexity levels. On the surface, this table is similar to most published tables comparing developmental levels across knowledge domains (Commons et al., 1990; Commons, Richards, & Armon, 1984). Such tables display theoretically analogous developmental levels from different knowledge domains on the same horizontal axis. The difference here is that the levels are not merely theoretically analogous. They are the same, because the same scoring criteria have been employed to score the texts from each domain. In other words, the Hierarchical Complexity Scoring System functions like a ruler, in the sense that it performs the same function regardless of context. This makes it possible to conduct direct cross-domain comparisons of conceptual development and to examine the relations between conceptions in different domains. For example, the abstract systems notion that *what we learn can be biased by existing knowledge* is implicit in the notion, also expressed for the first time at the abstract systems level, that *different interpretations of knowledge can result in different truths*.

________________insert Table 3 about here________________

Another insight offered by the juxtaposition of learning, truth, and moral concepts is that, overall, the learning and truth concepts appear to overlap more with one another than they do with moral concepts, suggesting that they actually represent the same knowledge domain or closely related knowledge domains. Finally, by juxtaposing conceptual development across knowledge domains or thematic strands, we can begin to understand how conceptions from one domain or strand become part of new conceptions on another domain or strand. For example, the single abstractions level conception of *learning as play* coordinates representational systems level conceptions of *learning* and *play* that, before the single abstractions level, were located on different thematic strands (*play* and *learning*) (Dawson, 1998). In the physics domain, we have similarly found that the notion that energy can exist in at least two states (conceptions of energy), along with an understanding that things can change from one form to another conceptions of change), are prerequisites to developing a concept of energy transformation. Insights like these are useful to educators who wish to promote the development of particular conceptualizations in that they provide us with insights into their conceptual prerequisites—what students need to learn before they can produce optimal constructions at the next level.

By examining patterns in the emergence of conceptions across complexity levels, Dawson and her colleagues (Dawson, manuscript submitted for publication; Dawson & Gabrielian, 2003) have demonstrated that it is possible to trace the development of individual concepts across numerous complexity levels, specifying how the conceptions of an earlier level are integrated to form the conceptions of a subsequent level. For example, the *learning as play* construct is further elaborated at the abstract mappings level when the relatively undifferentiated play component is replaced (or supplemented) with the concept of *interest*. Learning is now understood as an activity that can be interesting or engaging in and of itself, and therefore, much like play.

The descriptions in Table 3 have been compared to the stage definitions of other researchers (Dawson, in press-a, manuscript submitted for publication; Dawson & Gabrielian, 2003). In each case, there is great overlap between analogous stage definitions of other researchers and the descriptions of reasoning my colleagues and I have constructed by integrating independent hierarchical complexity and conceptual analyses. However, some significant differences have emerged. First, the descriptions of conceptual content developed employing our method are more extensive than those provided by other researchers. Second, because these are descriptions rather than definitions, they are open-ended. It is possible to add to these descriptions as subsequent samples reveal new conceptions at the various complexity levels, supporting cross-gender, cross-cultural, and cross-context comparisons of reasoning. Third, research has shown that these descriptions can be developed from cross-sectional rather than longitudinal data, meaning that conceptual development in a given domain can be described without the arduous and lengthy process of collecting longitudinal data, although confirmatory longitudinal evidence is desirable (Dawson, 2003a; Dawson, 2003d; Dawson & Gabrielian, 2003).

Research to Practice

Thus far, I have primarily discussed the pure research applications of developmental maieutics. Another important aspect of the methodology is that it ties together fundamental research and practice. Detailed accounts of conceptual development provide a valuable insight into how learning takes place. First, these accounts expose the sequence in which conceptual understandings are constructed and permit us to identify the conceptual precursors of concepts expressed at any given complexity level. For example, in an ongoing study of the development of leadership concepts conducted to the National Security Agency (NSA), we have found that NSA managers do not adequately coordinate the interests of employees and the organization until they can demonstrate elaborated conceptions of both (Dawson, 2003c). This coordination requires that both employees and organizations are conceptualized as complex entities with both shared and divergent interests that vary over time. Interestingly, we found that managers who demonstrated an elaborated conception of their organization, but lacked an elaborated conception of their employees, privileged the organization when they made management decisions. For example, when responding to a dilemma involving a female employee who wanted to cut back on work hours to care for her infant daughter, managers with less elaborated conceptions of the employee often advocated transferring or firing the employee in order to assure the success of their organization's mission. Conversely, managers who demonstrated an elaborated conception of their employees, but lacked an elaborated conception of their organization, privileged employees when they made management decisions. In this case, managers were willing to minimize the importance of the mission of their organization in order to provide support for the employee. Managers with elaborated conceptions of both the employee and the organization devised solutions that coordinated the needs of the new mother with the mission of the organization. As one manager explained, "I think the only way people are truly productive is if we [can] accommodate their personal needs along with the mission needs" (0108). Findings like those described here provide insights that are not only valuable from a diagnostic point of view, but can provide part of the basis for developing individualized learning programs, targeted to the needs of particular managers. In this instance, for example, a manager with an inadequately elaborated conception of the employee might benefit from engaging in learning activities that focus on promoting the development of a more elaborated conception of the employee prior to engaging in learning interventions that promote the coordination of the interests of employees with the interests of the organization.

Second, our research has shown that movement from one complexity level to the next depends, in a general sense, on the amount of conceptual elaboration that has occurred at the earlier complexity level (Dawson & Wilson, in press). This means that instructional interventions that are intended to move individuals to the next complexity level (restructuring) are likely to be more effective when conceptions at the current developmental level have become highly elaborated (learning). Conversely, when conceptions at the current complexity level are unelaborated, providing additional conceptual understandings at that level (learning) may be a more effective way to enhance development than attempting to promote movement to the next developmental level (structuring). We are presently testing this hypothesis in an intervention study of leadership development.

Third, the descriptions of conceptualizations associated with each complexity level can be translated into scoring rubrics for teachers, making it possible for persons untrained in hierarchical complexity theory to benefit directly from research results. One of the goals of our current investigation into the development of science concepts is the production of detailed scoring rubrics designed to help teachers track the developmental progress of their students, and to help them respond to students' developmental needs. In 2004, we will be evaluating teachers' use of such rubrics in a physical science program.

Fourth, linguistic analyses of the numerous texts we have gathered and scored over the last several years have revealed strong relations between complexity level and certain lexical and syntactic features of language. These analyses have helped to refine our ability to identify transitions from one complexity level to the next. By comparing linguistic features of transitional and consolidated performances, we are developing criteria for determining when a performance shows early evidence of the structures of a new complexity level. In fact, we have been able to employ what we have learned about the relation between language and complexity level to develop an accurate and reliable computer assisted developmental scoring system (Dawson & Wilson, in press). This system not only scores performances for complexity level, but also employs objective criteria to rank them within complexity levels. This automated system, called the Lexical Abstraction Assessment System (LAAS), makes it possible, for the first time, to conduct large-scale developmental assessments of texts. The implications for program evaluation, individual assessment, and research, are too numerous to consider here.

Discussion

In this chapter I have described a method for examining conceptual development that separates the measurement of structure from the assessment of content, thus making it possible to pose questions about their interrelations. I began by describing an alternative conception of learning as one of two interrelated processes in development. This was followed with a discussion of the principles of measurement, which are of central importance to this work because the hierarchical complexity scale is employed as a developmental ruler. I then described the hierarchical complexity construct in some detail and examined the validity and reliability of the Hierarchical Complexity Scoring System in terms of its validity as a measure. Following this, I discussed the relationship between complexity level and meaning and discussed how what we have learned about this relationship is being (or can be) applied in educational contexts.

Although no single methodological perspective can provide a complete account of cognitive development, the perspective offered here promises to engender important insights into adult cognitive development. While it clearly builds on the methods of earlier cognitive developmentalists, it avoids some of the limitations posed by these earlier methods, which, in failing to distinguish adequately between structure and content, limited the generalizability and applicability of research findings. A reliable and valid developmental ruler that can be applied in a wide range of contexts and content domains makes it possible to make meaningful inter-individual, cross-domain, cross-context, cross-gender, and cross-cultural comparisons of developmental progress. Moreover, this method can easily be employed to inform practice, and tools that have emerged from the research (scoring rubrics & the LAAS) can readily be employed in applied settings. In the past, valid and practical developmental assessments were difficult to implement because the methods for gathering and evaluating developmental data have been arduous and expensive. In this respect, the LAAS, our computerized developmental assessment system, makes large scale or frequent developmental assessments practical for the first time. Tools like the LAAS may well contribute to the transformation of developmental science, making it possible to address questions about adult development that were impractical to address with earlier assessment tools.

Acknowledgements

Special thanks to Cheryl Armon, Marvin Berkowitz, Larry Walker, the Murray Research Center and John Wise who have provided data employed in my analyses of adult conceptual development. The research reported in this paper was made possible in part by a grant from the Spencer Foundation. The data presented, the statements made, and the views expressed are solely the responsibility of the author.

References

Amir, R., Frankl, D. R., & Tamir, R. (1987). Justifications of answers to multiple choice items as a means for identifying misconceptions. In J. D. Novak (Ed.), *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics* (Vol. Vol. I, pp. 15-26). Ithaca, NY: Cornell University.

Armon, C. (1984). *Ideals of the good life: Evaluative reasoning in children and adults.* Unpublished Doctoral dissertation, Harvard University, Boston.

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

Armon, C., & Dawson, T. L. (2002). The good life: A longitudinal study of adult value reasoning. In J. Demick & C. Andreoletti (Eds.), *Handbook of Adult Development* (pp. 271-300). New York: Plenum Press.

Bandura, A. (1992). Social cognitive theory. In V. Ross (Ed.), *Six theories of child development: Revised formulations and current issues.* (pp. 1-60). London, England: Jessica Kingsley.

Brainerd, C. J. (1993). Cognitive development is abrupt (but not stage-like). *Monographs of the Society for Research in Child Development, 58*, 170-190.

Case, R., & Okamoto, Y. (1996). *The role of central conceptual structures in the development of children's thought* (Vol. 61 (1-2)). New York: Blackwell.

Case, R., Okamoto, Y., Henderson, B., & McKeough, A. (1993). Individual variability and consistency in cognitive development: New evidence for the existence of central conceptual structures. In R. Case & W. Edelstein (Eds.), *The new structuralism in cognitive development: Theory and research on individual pathways* (pp. 71-100). Basel, Switzerland: S. Karger.

Cavallo, A. M. L. (1996). Meaningful learning, reasoning ability, and students' understanding and problem solving of topics in genetics. *Journal of Research in Science Teaching, 33*(6.), 625-656.

Colby, A., Kohlberg, L., Gibbs, J., & Lieberman, M. (1983). *A longitudinal study of moral judgment* (Vol. 48 (1-sup-2)). New York.

Commons, M. L., Armon, C., Kohlberg, L., Richards, F. A., Grotzer, T. A., & Sinnott, J. (Eds.). (1990). *Adult development, Vol. 1: Comparisons and applications of developmental models*. New York: Praeger.

Commons, M. L., Richards, F. A., & Armon, C. (Eds.). (1984). *Beyond formal operations*. New York: Praeger.

Commons, M. L., Richards, F. A., with Ruf, F. J., Armstrong-Roche, M., & Bretzius, S. (1984). A general model of stage theory. In M. Commons, F. A. Richards & C. Armon (Eds.), *Beyond Formal Operations* (pp. 120-140). New York: Praeger.

Commons, M. L., Straughn, J., Meaney, M., Johnstone, J., Weaver, J. H., Lichtenbaum, E., et al. (1995, November). *The general stage scoring system: How to score anything.* Paper presented at the Annual meeting of the Association for Moral Education.

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

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

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

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

Dawson, T. L. (2002b, January). *Measuring intellectual development across the lifespan.* Paper presented at the Powerful learning & the Perry scheme: Exploring intellectual development's role in knowing, learning, and reasoning, California State University, Fullerton, CA.

Dawson, T. L. (2003a, June). *Assessing intellectual development: Three approaches, one sequence.* Paper presented at the Annual meeting of the Jean Piaget Society, Chicago.

Dawson, T. L. (2003b, 1/31/03). *The Hierarchical Complexity Scoring System*. Retrieved January, 2003, from http://gseacademic.harvard.edu/~hcs/

Dawson, T. L. (2003c, July). *Hierarchical complexity: A yardstick for developmental science.* Paper presented at the MBE Summer Institute, Harvard University.

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

Dawson, T. L. (in press-a). Assessing intellectual development: Three approaches, one sequence. *Journal of Adult Development.*

Dawson, T. L. (in press-b). A stage is a stage is a stage: A direct comparison of two scoring systems. *Journal of Genetic Psychology*.

Dawson, T. L. (manuscript submitted for publication). *"A good education is…" The development of evaluative thought across the life-span*.

Dawson, T. L., Commons, M. L., & Wilson, M. (in review). The shape of development.

Dawson, T. L., Commons, M. L., & Wilson, M. (manuscript submitted for publication). The shape of development.

Dawson, T. L., & Gabrielian, S. (2003). Developing conceptions of authority and contract across the life-span: Two perspectives. *Developmental Review, 23*, 162-218.

Dawson, T. L., & Wilson, M. (in press). The LAAS: A computerized developmental scoring system for small- and large-scale assessments. *Educational Assessment*.

Dawson, T. L., Xie, Y., & Wilson, M. (2003). Domain-general and domain-specific developmental assessments: Do they measure the same thing? *Cognitive Development*, 61-78.

Feldman, D. H. (1995). Learning and development in a nonuniversal theory. *Human Development, 38*, 315-321.

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

Fischer, K. W., & Ayoub, C. (1994). Affective splitting and dissociation in normal and maltreated children: Developmental pathways for self in relationships. In D. Cicchetti & S. L. Toth (Eds.), *Disorders and dysfunctions of the self: Rochester*

Symposium on Developmental Psychopathology, Vol. 5. (pp. 149-222). Rochester, NY: University of Rochester Press.

Fischer, K. W., Ayoub, C., Singh, I., Noam, G., Maraganore, A., & Raya, P. (1997). Psychopathology as adaptive development along distinctive pathways. *Development & Psychopathology, 9*.

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* (5 ed., pp. 467-561). New York: Wiley & Sons.

Fischer, K. W., & Granott, N. (1995). Beyond one-dimensional change: Parallel, concurrent, socially distributed processes in learning and development. *Human Development, 38*, 302-314.

Germann. (1994). Testing a model of science process skills acquisition: An interaction with parents' education, preferred language, gender, science attitude, cognitive development, academic ability, and biology knowledge. *Journal of Research in Science Teaching, 31*(7), 749-783.

Gilligan, C. (1977). In a different voice: Women's conceptions of self and of morality. *Harvard Educational Review, 47*, 481-517.

Halford, G. S. (1999). The properties of representations used in higher cognitive processes: Developmental implications. In I. E. Sigel (Ed.), *Development of mental representation: Theories and applications.* (pp. 147-168). Mahwah, NJ, USA: Erlbaum.

Inhelder, B., Bovet, M., & Sinclair, H. (1967). Development and Learning. *Psychologie V Ekonomicke Praxi, 26*(1), 1-23.

Kitchener, K. S., & King, P. M. (1990). The reflective judgment model: ten years of research. In M. L. Commons, C. Armon, L. Kohlberg, F. A. Richards, T. A. Grotzer & J. D. Sinnott (Eds.), *Adult development* (Vol. 2, pp. 62-78). New York: Praeger.

Kitchener, K. S., Lynch, C. L., Fischer, K. W., & Wood, P. K. (1993). Developmental range of reflective judgment: The effect of contextual support and practice on developmental stage. *Developmental Psychology, 29*, 893-906.

Lawson, A. E., Alkhoury, S., Benford, R., Clark, B. R., & Falconer, K. A. (2000). What kinds of scientific concepts exist? Concept construction and intellectual development in college biology. *journal of research in science teaching, 37*(9), 996-1018.

Lawson, A. E., & Renner, J. W. (1975). Relationships of concrete and formal operational science subject matter and the developmental level of the learner. *Journal of Research in Science Teaching, 12*(4), 347-358.

Lawson, A. E., & Thompson, L. D. (1988). Formal reasoning ability and misconceptions concerning genetics and natural selection. *Journal of Research in Science Teaching, 25*(9), 733-746.

Lawson, A. E., & Weser, J. (1990). The rejection of nonscientific beliefs about life: Effects of instruction and reasoning skills. *Journal of Research in Science Teaching, 27*(6), 589-606.

Lourenco, O., & Machado, A. (1996). In defense of Piaget's theory: A reply to 10 common criticisms. *Psychological Review, 103*, 143-164.

Mentkowski, M., Moeser, M., & Strait, M. J. (1983). *Using the Perry scheme of intellectual and ethical development as a college outcomes measure: a process and criteria for judging student performance, vol. 1*. Milwaukee, WI.: Alverno College, Office of Research and Evaluation.

Overton, W. (1998). Developmental psychology: Philosophy, concepts, and methodology. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of human development* (5 ed., pp. 107-188). New York: John Wiley & Sons.

Overton, W. F., Ward, S. L., Noveck, I. A., & Black, J. (1987). Form and content in the development of deductive reasoning. *Developmental Psychology, 23*.

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

Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development* (T. Brown & K. J. Thampy, Trans.). Chicago: The University of Chicago Press.

Piaget, J., & Garcia, R. (1989). *Psychogenesis and the history of science* (H. Feider, Trans.). New York: Columbia University Press.

Renner, J. W., & Marek, E. A. (1990). An educational theory base for science teaching. *Journal of Research in Science Teaching, 27*(3), 241-246.

Rose, S. P., & Fischer, K. W. (1989). *Constructing task sequences: A structured approach to skill theory* (Instructional Manual). Cambridge, MA: Harvard.

Shayer, M., & Adey, P. S. (1993). Accelerating the development of formal thinking in middle and high school students: IV. Three years after a two-year intervention. *Journal of Research in Science Teaching, 30*, 351-366.

Shweder, R. A. (1996). True ethnography: The lore, the law, and the lure. In *Ethnography and human development: Context and meaning in social inquiry.* (pp. 15-52). Chicago, IL, USA: The University of Chicago Press.

Shweder, R. A., Mahapatra, M., & Miller, J. G. (1987). Culture and moral development. In J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children* (pp. 1-83). Chicago: University of Chicago Press.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford.

Sinnott, J. D. (1994). The relationship of postformal thought, adult learning, and lifespan development. In J. Sinnott (Ed.), *Interdisciplinary handbook of adult lifespan learning.* (pp. 105-119). Westport, CT: Greenwood Press.

Smith, L. (1998). Learning and the development of knowledge. *Archives de Psychologie, 66*(258), 201-219.

Tappan, M. B. (1990). Hermeneutics and moral development: Interpreting narrative representations of moral experience. *Developmental Review, 10*, 239-265.

Walker, L. J., & Taylor, J. H. (1991). Stage transitions in moral reasoning: A longitudinal study of developmental processes. *Developmental Psychology, 27*, 330-337.

Wilson, M., & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education, 13*(2), 181-208.

---

[i] D*evelopmental maieutics,* a methodological perspective that has been strongly influenced by Overton's (1998) call for relational methods in developmental science, calls for an explicit effort to expose patterns of conceptual development in texts by isolating the hierarchical complexity dimension and then examining the interaction between this dimension and other features of the same texts. These other features include propositional content, thematic content, lexical content, and cross-domain or cross-context characteristics. The perspective further calls for an integration of fundamental research with practice by employing the hierarchical complexity construct and research results to inform curriculum design, assessment, and program evaluation.

[ii] There are also those who view learning (taking in knowledge) as the primary process in development (Bandura, 1992; Shweder, Mahapatra, & Miller, 1987).

[iii] At the single abstractions order, students commonly constructed a notion of energy as something "behind" the movement of objects, whereas students functioning at the representational systems order viewed energy and movement as the same thing.

[iv] When stages are defined in terms of particular conceptual content, it becomes possible to argue that (1) an individual is functioning at a given developmental level because he or she is capable of producing a particular conception, *and* that (2) an individual is capable of producing a particular conception because he or she is functioning at a particular developmental level.

[v] Single abstractions dominate by age 10 or 11 in most populations that have been sampled by developmental researchers. However, a small percentage of adults do not move beyond this complexity level in their moral reasoning.

Table 1: Orders of hierarchical complexity

Complexity order name	**Logical structure**	**Hierarchical order of abstraction (conceptual structure)**
Single reflexive actions	Single reflexes	1st order reflexes
Reflexive mappings	Coordinates one aspect of two or more reflexes	2nd order reflexes, which coordinate 1st order reflexes
Reflexive systems	Coordinates multiple aspects of reflexes	3rd order reflexes, which coordinate 2nd order reflexes
Single sensorimotor schemes	Abstracts single sensorimotor schemes from multiple reflexive systems	1st order sensorimotor schemes, which coordinate 3rd order reflexes
Sensorimotor mappings	Coordinates one aspect of two or more sensorimotor schemes	2nd order sensorimotor schemes, which coordinate 1st order sensorimotor schemes
Sensorimotor systems	Coordinates multiple aspects of sensorimotor schemes	3rd order sensorimotor schemes, which coordinate 2nd order

Single representations	Abstracts single representations from multiple sensorimotor schemes, often identifies one aspect of a single representation	1^{st} order representations, which coordinate 3^{rd} order sensorimotor schemes
Representational mappings	Coordinates one aspect of two or more representations	2^{nd} order representations, which coordinate 1^{st} order representations
Representational systems	Coordinates multiple aspects of representations	3^{rd} order representations, which coordinate 2^{nd} order representations
Single abstractions	Abstracts single abstractions from multiple representational systems, often identifies one aspect of a single abstraction	1^{st} order abstractions, which coordinate 3^{rd} order representations
Abstract mappings	Coordinates one aspect of two or more abstractions	2^{nd} order abstractions, which coordinate 1^{st} order abstractions
Abstract systems	Coordinates multiple aspects of abstractions	3^{rd} order abstractions, which coordinate 2^{nd} order abstractions
Single	Abstracts single principles from multiple abstract	1^{st} order principles, which coordinate 3^{rd} order

principles/axioms	systems, often identifies one aspect of a single principle	abstractions

Table 2: Education as preparation for the good life (Dawson, 1998)

A good education…	Number and percent of occurrences of each concept code at each complexity order (1 per case)			
	Single abstractions (n = 7)	**Abstract mappings (n = 38)**	**Abstract systems (n = 65)**	**Single principles (n = 38)**
provides life/ work skills/ knowledge.	1 (14.3%)	24 (63.2%)	28 (43.1%)	15 (39.5%)
prepares you for a work/ a job.	4 (57.1%)	17 (44.7%)	12 (18.5%)	5 (13.2%)
teaches survival skills.		3 (7.9%)	4 (6.2%)	4 (10.5%)
provides interpersonal/ social skills.		7 (18.4%)	14 (21.5%)	14 (36.8%)
produces better/ respectable/ responsible people		6 (15.8%)	14 (21.5%)	15 (39.5%)
prepares you achieve personal goals.		5 (13.2%)	6 (9.2%)	3 (7.9%)
prepares you to make money/ a living.		6 (15.8%)	8 (9.2%)	1 (2.6%)
prepares students to succeed.		5 (13.2%)	5 (7.7%)	2 (5.3%)
prepares students to contribute.		4 (10.5%)	5 (7.7%)	12 (31.6%)
enhances/ enriches/ betters your life.		7 (18.4%)	15 (23.1%)	8 (21.1%)
enhances self worth/ esteem/confidence/respect.			8 (9.2%)	3 (7.9%)
increases choices/ opens doors.			8 (9.2%)	
should round out/ broaden/ balance the student.			16 (24.6%)	7 (18.4%)
produces effective/ competent/ self-reliant people			7 (7.7%)	7 (18.4%)
prepares students to function in society.			10 (15.4%)	10 (26.3%)
leads to personal fulfillment/ meaning.			12 (18.5%)	8 (21.1%)
increases gratification/ satisfaction.			11 (16.9%)	5 (13.2%)
improves/ advances society.			12 (18.5%)	13 (34.2%)
promotes social/ global justice/ welfare.				8 (21.1%)
promotes autonomy/ self-				7

actualization/ self examination.	(18.4%)
produces good/ productive citizens.	6 (15.8%)
is empowering.	7 (18.4%)

Table 3: Summary of conceptions of learning, truth, and contract associated with the abstract mappings, abstract systems, and single principles complexity orders

Complexity order	**Learning** (Adapted from Dawson, in press; Dawson, manuscript submitted for publication)	**Truth** (Adapted from Dawson, in press)	**Contract/Promise** (Adapted from Dawson & Gabrielian, 2003)
Single abstractions	• Helps you understand things • Helps you remember, know things. • Provides skills for life. • Is not just from books (can be from projects, field trips, etc.). • Should be fun, interesting.	• Is findable. • Can be observed. (A broken window is a broken window.) • Is something you "just know." • Facts and opinions are different. Facts are truths.	• Keeping a promise means keeping "your word." A promise is a promise. • To assure another person that you will definitely do what you have agreed to do. • Keeping promises is important to

			friendship or trust; may cause others to keep promises to you. • Breaking promises makes you feel bad or guilty; makes your kids break promises. • It is okay to break a promise if you will never see the promisee again.
Abstract mappings	• Promotes understanding rather than memorization. • Involves taking in other people's opinions and 'mixing' them with your own. • Involves hands-on experience, learning from experience, learning from life in general.	• Is fact, factual knowledge, proven knowledge. • Something to figure out, but hard to find. • Found through problem-solving. • Can be proven (if science-based). • There are some absolute scientific truths, scientific laws.	• Are agreements that have a sacred quality/should not be made lightly. • Build trust, which you must have to maintain friendships. • Should be kept because you would want others to do the same for you; because breaking them can cause others to lose respect for you or

	• Happens over time. • Provides a broad base of knowledge. • Can be applied in real life.	• Some scientific truths can change, be altered. • What is accepted as truth may not be real truth. • Individuals have their own personal truths, based on upbringing, opinions. • People can fail to understand truth for psychological reasons (need to make self look good, bias, upbringing). • Universal truths are fixed, external. • Universal truths are taken on faith (esp. if associated with religion).	ruin your reputation; or cause psychological harm. • Whether or not one should keep a promise can depend upon its importance.
Abstract	• Involves 'processing'	• Is objective.	• Set up binding rules of behavior

systems	knowledge • Involves interpretation. • Involves being exposed to different perspectives. • Involves learning how to learn • Can be biased by existing knowledge. • Develops reasoning skills. • Promotes development. • Promotes self-understanding, understanding of others.	• Is bound to interpretation. • There is no absolute truth. • There are certain overall truths. • Pursuit of truth is an ongoing process. • Some methods of seeking truth are better than others. • Interpretation of knowledge can differ, resulting in different 'truths' • Can be more absolute or concrete in some knowledge domains than in others. • Can't be entirely relative. • Can be multiple truths in some things.	for the individuals involved in making them. • Have the purpose of maintaining relationships by building trust; make it possible for people to live together. • Should be kept because the actions of others are influenced by a promise; to preserve your own or others' perceptions of your integrity. • Because of the nature of a promise, it is not okay to break one with no explanation.

		• Scientific truth and religious truth should be/are evaluated differently.	
Single principles	• Involves restructuring existing knowledge. • A thought process that is partially conscious and partially subconscious, a network integrating new ideas. • Involves seeking a principle that holds together a perspective. • Develops a critical perspective. • Promotes critical discourse.	• Is reevaluated throughout our lives. • Human desire is to seek out truth. • Two separate truths can be interrelated to form a third truth. • Some mathematical truths may be absolute. • Universal truths are things all reasonable persons can agree about at a given moment in time.	• Are mutual obligations set up by persons acting as autonomous agents; are uniquely human. • Are the basis for the mutual trust that makes human society possible. • Without the expectation that promises will be kept, contracts honored, society could not exist. • One should only break a promise to serve a higher ideal.

References

Dawson, T. L. (in press). Assessing intellectual development: Three approaches, one sequence. *Journal of Adult Development*.

Dawson, T. L. (manuscript submitted for publication). *"A good education is…" The development of evaluative thought across the life-span*.

Dawson, T. L., & Gabrielian, S. (2003). Developing conceptions of authority and contract across the life-span: Two perspectives. *Developmental Review, 23*, 162-218.

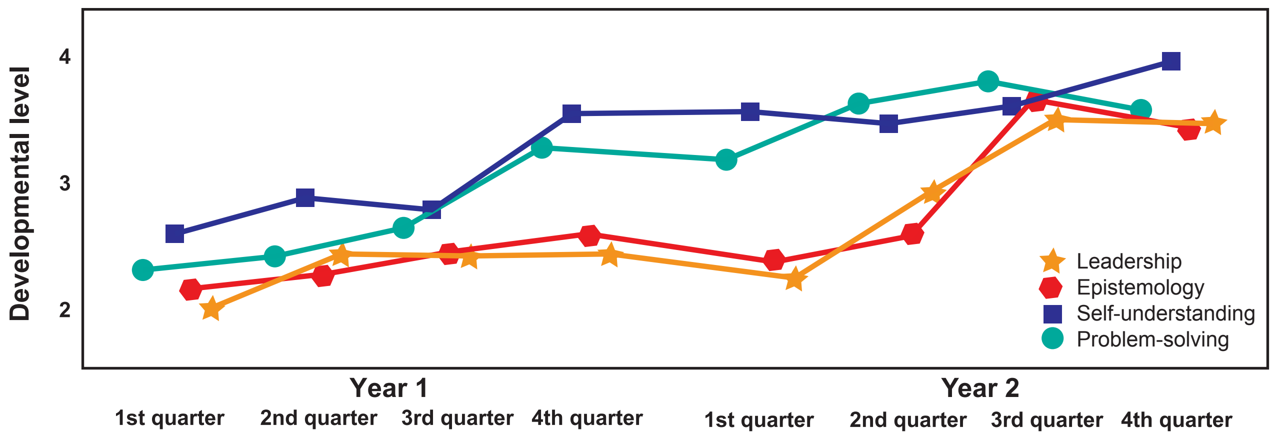

Figure 1: Lucile's developmental progress in four knowledge domains

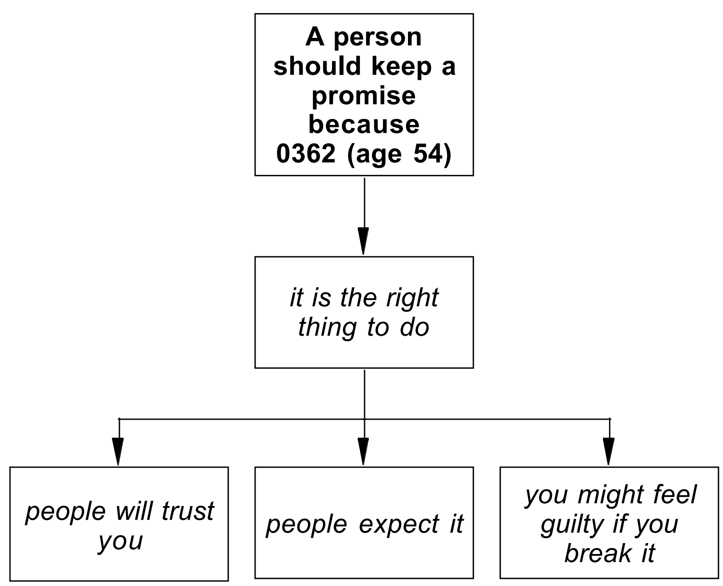

Figure 2: Conception of promise, single abstractions order

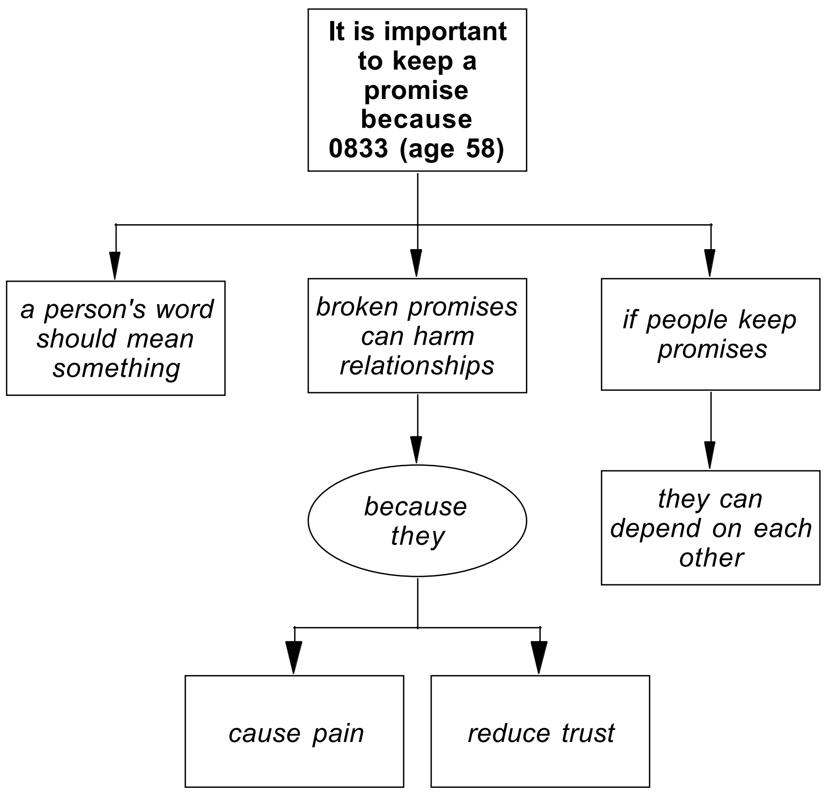

Figure 3: Conception of promise, abstract mappings order

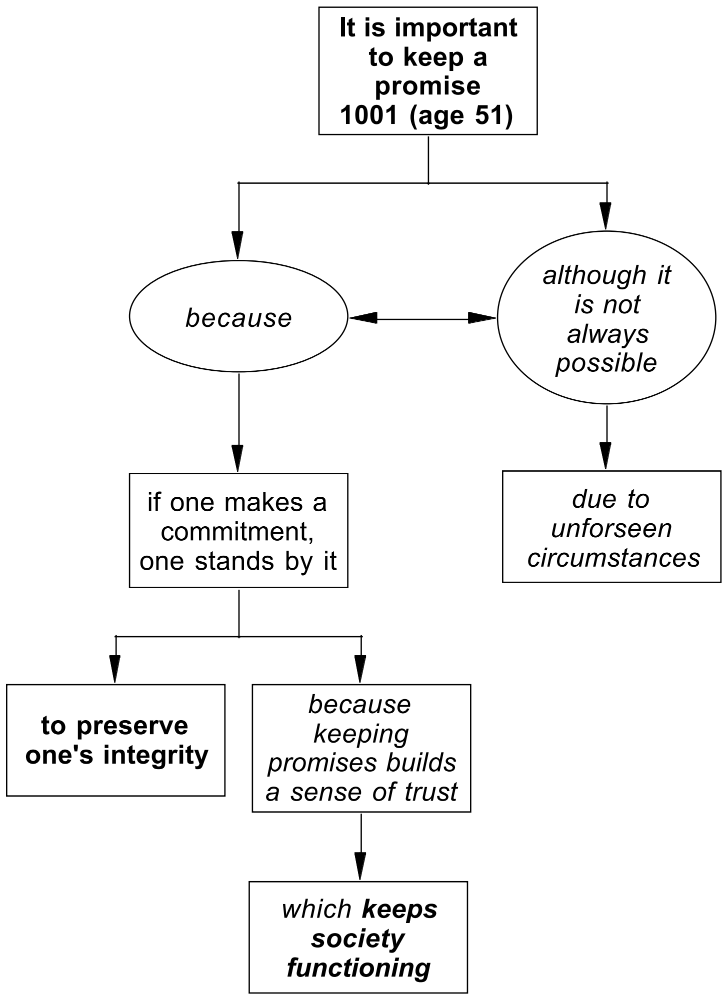

Figure 4: Conception of promise, abstract systems order

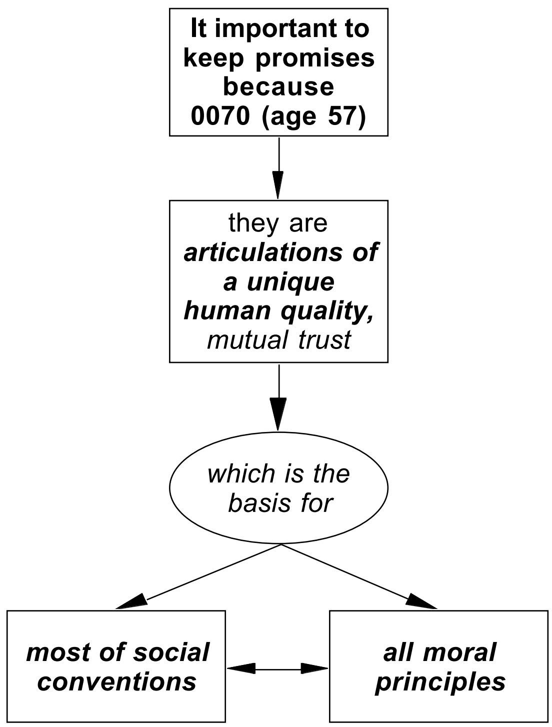

Figure 5: Conception of promise, single principles order