BEHAVIORAL DEVELOPMENT BULLETIN

VOLUME 19
NUMBER 4
DECEMBER 2014

There is only one stage domain

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

1 Dare Institute
2 Harvard Medical School
3 Harvard University

ABSTRACT

The study used the model of hierarchical complexity (MHC) to test the theory that different domains in development would develop in synchrony, allowing an individual to solve tasks from various domains using the same mental structure for each task. The MHC instruments used were the empathy, helper person, counselor patient, breakup, caregiver, algebra, balance beam, infinity and laundry instruments. The instruments can be categorized as belonging to two different subdomains, the social subdomain, and the logic/mathematics/physical sciences subdomain. Instruments in this social subdomain measure developmental stage in a variety of social contexts. These social contexts included empathy for person after an accident, guidance and assistance by a helper, counseling patients, understanding romantic breakups, and caring for children and infants. The other subdomain is composed of mathematical (algebra & infinity), logical (laundry), and physical science (balance beam). In order to conclude how related the performances were, three analyses were carried out.

First, Rasch analysis yielded person scores akin to person stage scores. Second, regression analysis was conducted to assess how well the order of hierarchical complexity (OHC) of the items predicted the Rasch difficulty of the items. Third a principal axis factoring was carried out with the person Rasch scores for every instrument. Irrespective of domains, if each instrument loaded on the first factor with all the factor scores over .7 and if the first factor accounted for more than 70% of the variance, then that would show that all instruments were part of a single domain. In each case the MHC accounted for a large amount of variance with r values over .7.

The principal axis factoring showed that person scores on each instrument loaded on the first factor (90.51% of the variance). All the factor scores on the first factor were over .85. There were very low loadings only on the second factor (4.947% of the variance). This implies that the instruments from the social subdomain and instruments from the logic/mathematics/physical sciences belong to a single domain.

KEYWORDS: synchrony, one domain, model of hierarchical complexity

Gilligan (1992) claimed that moral judgment and care formed different domains. Perry's 1968 study, 'Forms of Intellectual and Ethical Development in the College Years' explored the possibility of two domains: epistemological and meta-ethical development. Erdynast, Armon and Nelson's 1978 study Cognitive-Developmental Conceptions of the True, the Good and the Beautiful researched across three domains: conceptions of factual reality, conceptions of the good, and conceptions of the beautiful. Erdynast and Chen's (2014) study on the 'Relations between Adult Developmental Conceptions of the Beautiful and Moral Development' claimed that there are four theoretically distinct domains. These four domains are: I) the real, II) the good, III) the just and right and IV) the beautiful.

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

Kegan (1979), Kohlberg, Levine & Hewer (1983) and Commons & Richards (2003) however, asserted that there was just one domain. This study used a factor analysis of nine very different stage instruments to see how many factors are found. If the factor scores of the nine instruments loaded highly on the first factor, i.e. the first factor accounted for greater than 70% of the variance, then that would mean that all instruments were part of a single domain. This would mean that task performances across subdomains form a single domain, supporting Inhelder and Piaget's (1958) theory of synchrony of development.

Inhelder & Piaget (1958) found synchrony in the logical, mathematical and physical sciences domains. For decades this theory has been studied. The findings have been largely in contrast with the theory of synchrony. Fischer (1980) points out that these studies not only showed that décalage, also known as asynchrony, was the norm, but that the amount of asynchrony was often great. These studies do not make clear how tight concurrence must be to support synchrony. They did not construct instruments to do the studies to show synchrony between domains.

» REFLECTIVE ABSTRACTION, HIERARCHICAL INTEGRATION, AND HIERARCHICAL COMPLEXITY

Many cognitive-developmental researchers assert that development in different knowledge domains does not necessarily proceed at the same rate (Fischer & Bidell, 1998; Lourenco & Machado, 1996). However, there is still considerable disagreement about whether development in different domains can be characterized in terms of a single generalized process. Domain theorists argue that different processes apply in different knowledge domains (Kohlberg, 1969; Larivee, Normandeau, & Parent, 2000; Turiel, 1980). Others, though they acknowledge that unique structures and processes are associated with particular domains also argue that a single general developmental process applies across domains (Case, Okamoto, Henderson, & McKeough, 1993; Fisher & Bidell, 1998). Piaget (2000) called this general process of reflective (or reflecting abstraction) abstraction, through which the actions of one developmental level become the subject of the actions of the subsequent level. The product of reflective abstraction is *hierarchical integration*. In conceptual development, hierarchical integration is observable in the concepts constructed at a new level by coordinating (or integrating) the conceptual elements of the prior level. These new concepts are said to be more hierarchically complex than the concepts of the previous level, in that they integrate earlier knowledge into a new form of knowledge. For example, independent conceptions of *play* and *learning* constructed at one complexity level are integrated into a conception of *learning as play* at the next complexity level (Dawson-Tunik, 2004a). Though it builds on the original *play* and *learning* concepts, the new concept cannot be reduced to the original *play* and *learning* elements. Not only is there a new concept in the recognition that learning can be playful, but the individual meanings of the elements *learning* and *play* have changed. The meanings of *learning* and *play* have changed in that each now incorporates some of the meaning embedded in the new construction. The concept of *learning* now includes *play* as a component, and the concept of *play* includes *learning* as a component.

A number of researchers have described developmental sequences that elaborate the basic notion of hierarchical integration, including Fischer (1980), who has emphasized the development of skill hierarchies in particular contexts. Commons and his colleagues (Commons et al., 1998) have described a task structure hierarchy. Pascual-Leone and Goodman (1979) have focused on the growth of mental attention and memory capacity. Case (1991) has described the development of memory capacity ad associated processing structures. Finally Demetriou and Valanides (1998) have described hierarchical development in terms of processing functions.

Not only are there definitional correspondences among analogous levels described by Commons, Fischer, and Piaget, there is empirical evidence of correspondences between complexity levels, skill levels, and orders of hierarchical complexity and at least three domain-based systems, including Kitchener and King's (Dawson, 2002b; King, Kitchener, Wood, & Davison, 1989; Kitchener & King, 1990; Kitchener, Lynch, Fischer, & Wood, 1993) stages of reflective judgment, Armon's good life stages (Dawson, 2002a), Perry's stages of epistemological development (Dawson, 2004), and Kohlberg's moral stages (Commons et al., 1989; Dawson & Gabrielian, 2003; Dawson, Xie, & Wilson, 2003). These correspondences suggest that, as a community, this group of developmental researchers is moving toward a consensus regarding the detection and aspects of the definition of developmental stages.

» THE MODEL OF HIERARCHICAL COMPLEXITY

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

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

Table 1. the 17 known orders of hierarchical complexity

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

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

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

» THE RASCH MODEL

Whereas they are well-known in psychometric circles, Rasch's (1980), models for measurement have been employed by developmental psychologists only recently (Andrich & Constable, 1984; Bond, 1994; Dawson, 1998, 2000, 2002; Draney, 1996; Muller, Sokol, & Overton, 1999; Wilson, 1984). These models are designed specifically to examine hierarchies of person and item performance, displaying both person proficiency and item difficulty estimates along a single interval scale (logit scale) under a probabilistic function. In addition, they can be employed to test the extent to which items or scores conform to a theoretically specified hierarchical sequence. A central tenet of stage theory is that cognitive abilities develop in a specified sequence, making the statistical tests implemented in a Rasch analysis especially relevant to understanding stage data. The Rasch model permits researchers to address questions like, "Are all single abstractions items more difficult than all representational systems level and less difficult than all abstract mappings items?" Moreover, the detailed information about item functioning and individual performances provided by the software makes it possible to simultaneously examine group and individual effects. These properties make Rasch models uniquely suitable for the investigation of many developmental phenomena.

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

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

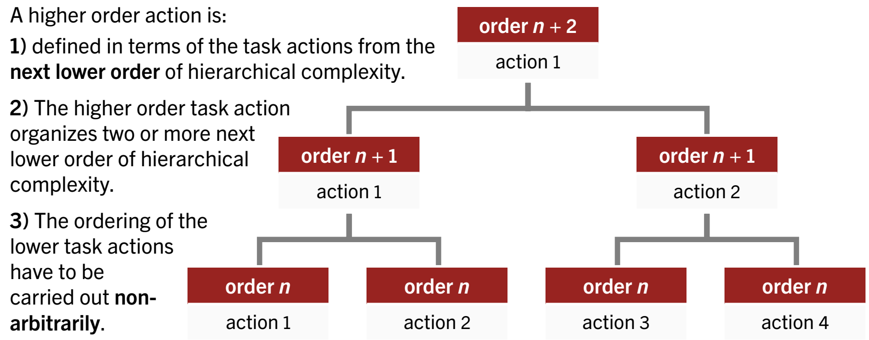


Figure 1. the three axioms of the MHC

physical sciences subdomain. The instruments were categorized as either belonging to the ***mathematical*** (Algebra & Infinity), ***logical*** (Laundry version of the Inhelder and Piaget's (1958) Pendulum problem), or ***physical science*** (balance beam) areas.

The second sets of instruments were from the social subdomain. Sequences of vignettes (series of similar stories following a common theme) from the Counseling, Child Caregiving Empathy, Helping, and Romantic Breakup sequences were used.

Sequences of vignettes used for Social Subdomain were Counseling, Child Caregiving Empathy, Helping, and Romantic Breakup. Each vignette sequence was followed by questions. For Laundry, Algebra, Infinity and Balance Beam instrument, sequence of mathematical and logic based problems were used. There were a total of 430 item represented in all the instruments together.

Each vignette sequence and problem sequence was at a different order of hierarchical complexity. These included:

1. Preoperational
2. Primary
3. Concrete
4. Abstract
5. Formal
6. Systematic
7. Metasystematic

» METHOD

Participants

There were a total of 409 participants among all instruments from two subdomains. The number participant for each instrument is shown in Table 2. The Age of the participants ranged from 12 years to 87 years. The mean age and standard deviation was $M = 28.42$ ($SD = 8.56$). Education level of the participants varied from elementary school graduate to graduate school; $M = 4$ ($SD = 1.27$). The mean of 4 represents a high school degree

Instruments

Two sets of instruments were used to assess participants' developmental stage. The first sets of instruments were from the logic/mathematics/

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

instrument	number of participants
helper person	40
counselor patient	30
empathy	22
breakup dilemma	21
caregiver	42
laundry	111
algebra	40
balance beam	51
infinity	52

Instrument description for social subdomain

1. Empathy: a person explains to another person how to understand the feelings of others in all the vignettes in this sequence
2. Counselor patient: a counselor explains treatment options to a patient in all the vignettes in this sequence
3. Breakup: a person reacts to breaking up with their partner in all the vignettes in this sequence
4. Helper person: a person gives another person guidance and assistance in all the vignettes in this sequence
5. Caregiver: in this sequence of vignettes a caregiver comforts a crying child in all the vignettes in this sequence
An example of social subdomain vignette sequence and corresponding questions is shown in Appendix J.

Instrument description for logic/mathematics/ physical sciences subdomain

1. ***Algebra (mathematics):*** this instrument sequence asked participants solve standard algebraic problems
2. ***Balance beam (physics):*** this instrument sequence was derived from inhelder and Piaget (1958) balance beam task
3. ***Infinity (mathematics):*** this instrument sequence asked participants to solve problems that dealt with conceptualizing infinity
4. ***Laundry (logic and mathematics):*** this instrument sequence was a version of Inhelder & Piaget (1958) pendulum task. It asked participants to detect causal relationships from various systems and then compare the systems using logic.

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

» PROCEDURE

Participants were asked to rate their preferences for vignette based questions for instruments from the social subdomain. 1–6 rating scale was used as answers to vignette based questions. For instruments from the logic/mathematics/physical sciences subdomain were asked to solve mathematical and logic based problems. The instruments from Social subdomain and logic/mathematics/ physical sciences subdomain were given in pairs. Different combination of instruments was used. There were a total of 430 item represented in all the instruments together

» RESULTS

There were three parts to the analysis. First, Rasch analysis of all the items from all the instruments was run. This is a psychometric analysis of the responses to the items showing their relative difficulty. The Rasch analysis confirmed that the items were performing as predicted by the MHC. Rasch figures are included in the Appendix so that one can see the actual raw Rasch scores for each item. The Rasch scores showed the raw magnitude of difficulty for each item. Through close inspection of the Rasch map, one can see which items are out of order.

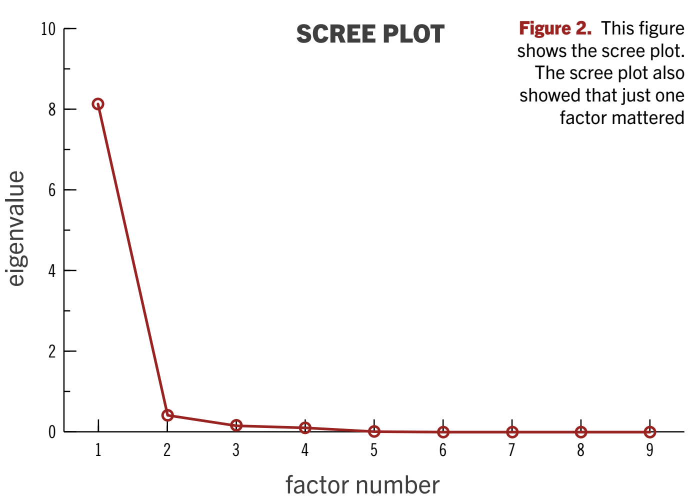


Figure 2. This figure shows the scree plot. The scree plot also showed that just one factor mattered

The Rasch maps in the appendix show that metasystematic items generally had the highest Rasch item scores (RIS) and preoperational items generally had the lowest Rasch item scores. This shows that metasystematic items were the most difficult and preoperational items the most easiest as predicted by MHC. There was some intermixing of items from adjacent stage. For example, in the counselor patient Rasch map, Primary stage items had a higher raw Rasch score in comparison to preoperational stage items. Similarly, formal stage items had a higher raw Rasch score in comparison to systematic stage items. In Infinity Rasch map, some systematic stage items had a higher raw Rasch scores in comparison to the metasystematic stage item. However, intermixing is not too terrible as shown by the high *r*'s for each instrument. The *r*'s for the instrument ranged from .711 to .980. Compared to other instruments, these are very high *r* values.

Table 3. principal axis factoring with the person Rasch scores for each instrument

		initial eigenvalues	
factor	total	% of variance	cumulative %
1	8.146	90.51	90.510
2	0.445	4.947	95.456
3	0.190	2.114	97.570
4	0.129	1.431	99.002
5	0.034	0.376	99.378
6	0.026	0.285	99.663
7	0.016	0.175	99.837
8	0.010	0.108	99.946
9	0.005	0.054	100.00

Second, we report regression analysis assessing how well the order of hierarchical complexity (OHC) of the items predicted the difficulty of the items in terms of their Rasch scores. Because the items have orders of hierarchical complexity (OHC), the relationship between the OHC and the Rasch item scores is psychophysical. There is an analytic property of a stimulus predicting a scaled response. A linear regression shows that order of hierarchical complexity (OHC) of the items predicts the actual difficulty of the items as represented by Rasch items scores. The linear regression showed that hierarchical complexity strongly predicted the stage of performance of all the items from all of the instruments; $r(430) = 0.959$, $p < 0.0001$.

Third, a principal axis factoring was carried out with the person Rasch scores for each instrument. Irrespective of domains, if each instrument loaded on the first factor with all the factor scores over .7 and if the first factor accounted for more than 70% of the variance, then that would show that all instruments were part of a single domain. The analysis showed that irrespective of domains, the person scores on each instrument loaded on the first factor (90.510% of the variance). There were only very low loadings on the second factor (4.947% of the variance) as shown in Table 2. The scree plot also showed that only the first factor mattered. This is shown in Figure 2. All the factor scores on the first factor were over .85 as shown in Table 4.

In each case the MHC accounted for the largest amount of variance with *r* values of $r(40) = .977$, $p \leq .001$ for helper person, $r(30) = .934$, $p \leq .001$) for counselor patient, $r(22) = .910$, $p \leq .001$) for empathy, $r(21) = .835$, $p \leq .001$) for breakup dilemma, $r(42) = .711$, $p \leq .001$ for caregiver, $r(51) = .980$, $p \leq .001$) for balance beam, $r(40) = .966$, $p \leq .001$) for algebra, $r(111) = .964$, $p \leq .001$) for laundry, and $r(52) = .912$, $p \leq .001$ for the infinity instrument. Rasch map and result of stepwise regression for all the instruments is shown in the appendix section.

The critical test for determining if there is just a single domain is to conduct a factor analysis. The first factor accounted for 90.510% of the variance. The second factor accounted for only

4.947% of the variance. This showed only one factor mattered. That shows all the instruments are part of a single domain

Table 4 shows the factor scores of all the instruments in the first factor and second factor. For missing values, cases were excluded listwise. All the factor scores on the first factor were over .85. Overall factor scores show the presence of just one factor. The second factor does not show in the factor matrix.

Table 4. factor matrix

instruments	factor 1	factor 2
helper person	0.973	0.129
infinity	0.968	0.142
counselor patient	0.964	–0.060
caregiver	0.962	–0.174
depression breakup	0.962	–0.181
algebra	0.952	0.093
empathy	0.942	–0.102
laundry	0.934	–0.271
balance beam	0.873	0.455

note: extraction method: principal axis factoring

» DISCUSSION

MHC assumes axiomatically that it is content and context free. A predicted consequence of this is that stage of performance on tasks of a given order of hierarchical complexity (OHC) should not depend on the content and context of the items in an assessment. This prediction was born out in the present study.

Here, order of hierarchical complexity (OHC) of the items almost perfectly predicted Rasch performance scores that reflected actual performance-measured difficulty among these instruments. The outcome was that the factor analysis found only a single factor. This implies that the instruments in social subdomain and instruments in the logic/mathematics/physical sciences subdomain belong to a single domain. This may explain why Inhelder & Piaget (1958) thought that development was synchronous across all tasks and domains.

Consider the two tasks presented next. They differ in content, form and method. The content and form, method of assessment of the helper-person task is entirely different from the content, form and method of assessment in the laundry task. The helper-person task has social content, a series of vignettes that are assessed by having the participants rate how well the method of interacting with the person to be helped (method). The laundry problem has chemical content. It presents how various ingratiated and washing methods might make a stain in cloth come out clean or dirty. They make predictions that are right or wrong. Yet they both load extremely highly on a single factor and the order of hierarchical complexity predicts the difficulty of performance extremely well.

What is the status of content and context claims about differences in stage given their lack of empirical support found in our study. Maybe what other views of domains reflect, is that content analysis is possible. One can analyze the content of a task and the method of presentation, but that does not mean that the content or method make a difference in stage. But we found content and form of the tasks to be unrelated to the stage of performance. That does not mean there are not people with performance deficits in given domains. These performance deficits may be because of lack of practice due to a lack of interest in the domain or lack of exposure to tasks in the domain or in some cases specific deficits results in unevenness in performance.

Content analysis has some use in describing the differences in the type of tasks to be completed. This does not mean that the task demands are identical or tasks actions are all the same. It just says that OHC required by different tasks is what is telling. In the small effects paper (Commons, Giri & Harrigan, 2014), we give further evidence that although there are many sources of differences, they are all overwhelmed by the predictive power of OHC.

We conclude there is only one stage domain. In the future, we plan to see what the relationship is between this domain and other measures of "smarts" such as IQ. ■

REFERENCES

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

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

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

Case, R. (1991). The mind's staircase: Exploring the conceptual underpinnings of children's thought and knowledge. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

Commons, M. L., Armon, C., Richards, F. A., Schrader, D. E., Farrell, E. W., Tappan, M. B., & Bauer, N. F. (1989). A multidomain study of adult development. In D. Sinnott, F. A. Richards, & C. Armon (Eds.), Adult development, Vol. 1: Comparisons and applications of developmental models. (pp. 33 – 56). New York, NY: Praeger

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

Commons, M.L., & Richards, F. A. (2003). Four Postformal Stages. In Demick, J. & Andreoletti, C. (Eds.). Handbook of adult development (pp. 199–219), New York: Plenum Press.

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

Commons, M. L., Giri, S. P., & Harrigan, W. H. (2014). The small effects of non-hierarchical complexity variables on performance. In Press.

Dawson, T. L. (1998). ''A good education is.'' A lifespan investigation of developmental and conceptual features of evaluative reasoning about education. Unpublished doctoral dissertation, University of California at Berkeley, Berkeley, CA.

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

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

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

Dawson, T. L. (2002c). New tools, new insights: Kohlberg's moral reasoning stages revisited. International Journal of Behavioral Development, 26, 154 – 166.

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

Dawson, T. L., Xie, Y., & Wilson, M. (2003). Domain-general and domain-specific developmental assessments: Do they measure the same thing? Cognitive Development, 18, 61 – 78.

Dawson-Tunik, T. L. (2004a). ''A good education is . . .'' The development of evaluative thought across the lifespan. Genetic, Social, and General Psychology Monographs, 130(1), 4 – 112.

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

Demetriou, A., & Valanides, N. (1998). A three-level theory of the developing mind: Basic principles and implications for instruction and assessment. In R. J. Sternberg & W. M. Williams (Eds.), Intelligence, instruction, and assessment: Theory into practice (pp. 211 – 246). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

Erdynast, A., Armon, C., & Nelson, J. (1978). Cognitive developmental conceptions of the true, the good, and the beautiful. The eighth annual proceedings of Piaget and the helping professions. Los Angeles: University of Southern California Press.

Erdynast, A., & Chen, W. (2014). Relations between Adult Developmental Conceptions of the Beautiful and Moral Development. In Press.

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

Fischer, K. W., & Bidell, T. R. (1998). Dynamic development of psychological structures in action and thought. In W. Damon & R. M. Lerner (Eds.), Handbook of child psychology: Theoretical models of human development (5th ed., pp. 467 – 561). New York, NY: Wiley.

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

Kegan, R. G. (1979). The Evolving Self: A Process Conception for Ego Psychology. The Counseling Psychologist, 8, 2, 5–34.

King, P. M., Kitchener, K. S., Wood, P. K., & Davison, M. L. (1989). Relationships across developmental domains: A longitudinal study of intellectual, moral, and ego development. In M. L. Commons, J. D. Sinnot, F. A. Richards, & C. Armon (Eds.), Adult development. Volume 1: Comparisons and applications of developmental models (pp. 57 – 71). New York, NY: Praeger.

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

Kohlberg, L., Levine, C., & Hewer, A. (1983). Moral stages: A current formulation and a response to critics. (pp. 63–128).Basel: Karger.

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. Goslin (Ed.), Handbook of socialization theory and research (pp. 347 – 480). Chicago, IL: Rand McNally.

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

Larivee, S., Normandeau, S., & Parent, S. (2000). The French connection: Contributions of French-language research in the post-Piagetian era. Child Development, 71(4), 823 – 839.

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

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

Pascual-Leone, J., & Goodman, D. (1979). Intelligence and experience: A neoPiagetian approach. Instructional Science, 8, 301 – 367.

Perry, W.G. (1968). Forms of intellectual and ethical development in the college years: A scheme. New York: Holt, Rinehart & Winston.

Piaget, J. (2000). Studies in reflecting abstraction (R. L. Campbell, Trans.). Hove, UK: Psychology Press.

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

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

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

Turiel, E. (1980). The development of social-conventional and moral concepts. In M.

Windmiller, N. Lambert, & E. Turiel (Eds.), Moral development and socialization (pp. 69 – 106). Boston, MA: Allyn & Bacon.

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

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

APPENDIX

» APPENDIX A1

Helper person rasch map

```
pre   = Preoperational
pri   = Primary
conc  = Concrete
abs   = Abstract
fl    = Formal
sys   = Systematic
me    = Metasystematic
```

```
                Higher Stage
 2                  +
                    |
                    |T
                    |
                 X  |  me1.1   me1.2
                 X  |  me1.3
                   T|  me1.4
                    |  me1.5   sys1.2
 1                  +S
                XX  |  f1.2    sys1.1
                 X  |  sys1.3  sys1.4  sys1.6
                   S|  f1.1    f1.4    sys1.5  sys1.7
                 X  |  f1.5
          XXXXXXXX  |  abs1.1  abs1.2  abs1.4  f1.3    f1.7
            XXXXXX  |  abs1.3  f1.6
               XXX M|M abs1.6  abs1.7
 0                  +  abs1.5
              XXXX  |
                XX  |
                 X  |
            XXXXXX S|  pri1.1
                 X  |S pri1.2
                 X  |  pri1.3  pri1.4
                    |  pre1.1  pre1.3  pri1.5  pri1.6  pri1.7
-1               X T+  pre1.2  pre1.4  pre1.5  pre1.6  pre1.7
                 X  |
                    |
                    |
                    |T
                    |
                    |
                    |
-2                  +
               Lower Stage
```

There were 40 participants ranging in age from 19 to 63 ($M = 22.16$, $SD = 9.09$) with education varying from elementary school to graduate school. Each × in the Rasch map represents a participant. In this figure, the concrete stage items are missing because; the items did not perform as designed.

» APPENDIX B1

Counselor patient rasch map

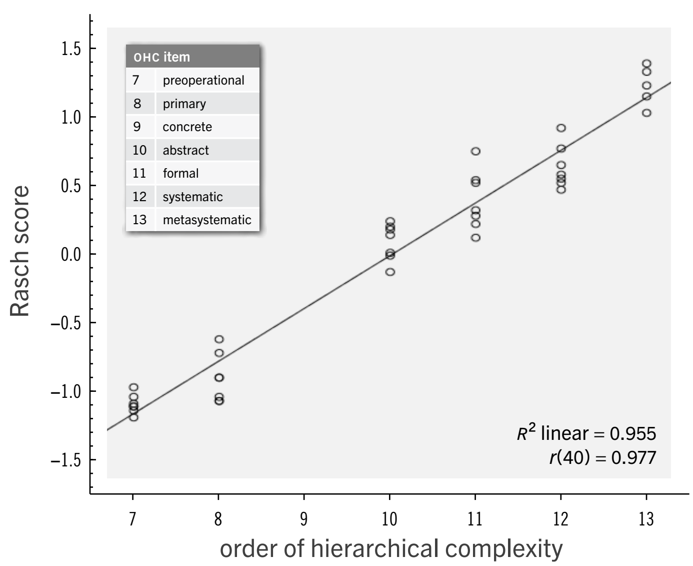
```
pre   = Preoperational
pri   = Primary
conc  = Concrete
abs   = Abstract
fl    = Formal
sys   = Systematic
me    = Metasystematic
```

```
                Higher Stage
 2                  +T
                    |
                 X  |  me1.3
                    |  me1.2
                   T|  me1.1   me1.4   me1.5   me1.6
                    |  me1.7
                    |  f11.1   f11.2   f11.6
 1                  +S f11.3   f11.4
                XX  |  f11.5
               XXX S|  f11.7
                 X  |  sys1.1  sys1.2
               XXX  |  sys1.4  sys1.5  sys1.6
                    |  sys1.3  sys1.7
                    |
 0              XX  +
               XXX M|M abs1.6
             XXXXX  |  abs1.1  abs1.2  abs1.3  abs1.5  abs1.7
                XX  |  abs1.4
             XXXXX  |  cr1.3   cr1.4   cr1.6
                    |  cr1.1   cr1.2   cr1.5
                   S|  cr1.7
-1                  +
                    |  pre1.4
                    |S pre1.1  pre1.2  pre1.6
                 X  |  pre1.3  pre1.5  pre1.7
                XX T|  pri1.2  pri1.4  pri1.5
                    |  pri1.1  pri1.3  pri1.6  pri1.7
                    |
-2                  +
                Lower Stage
```

There were 30 participants ranging in age from 16 to 71 ($M = 25.13$, $SD = 7.02$) with education varying from high school to graduate school. Each × in the Rasch map represents a participant

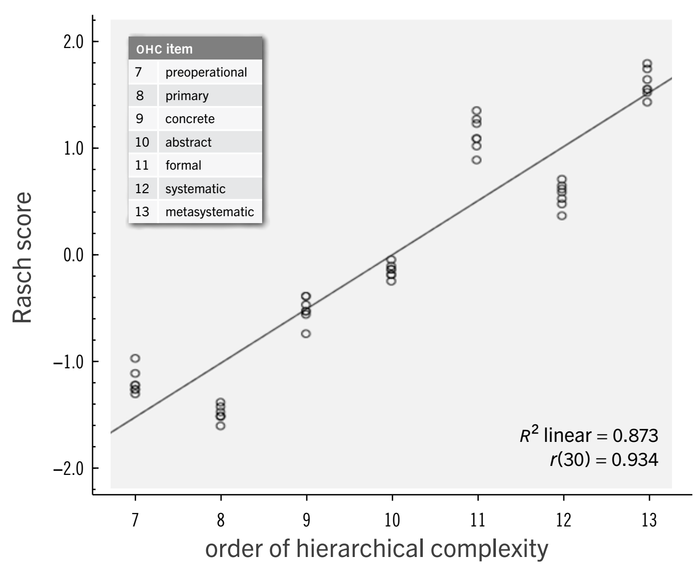
» APPENDIX A2

Helper person results

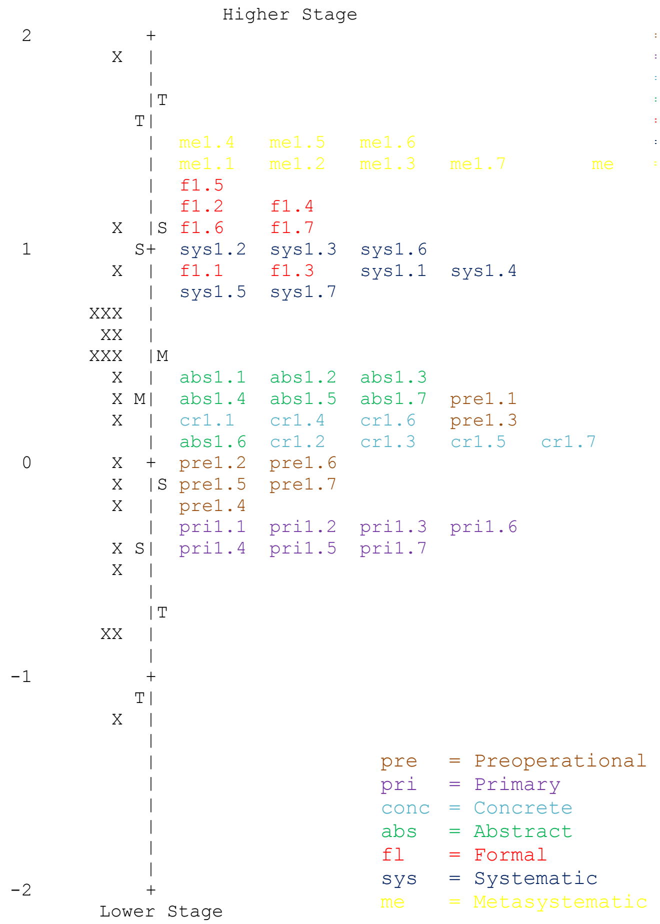


» APPENDIX B2

Counselor patient results

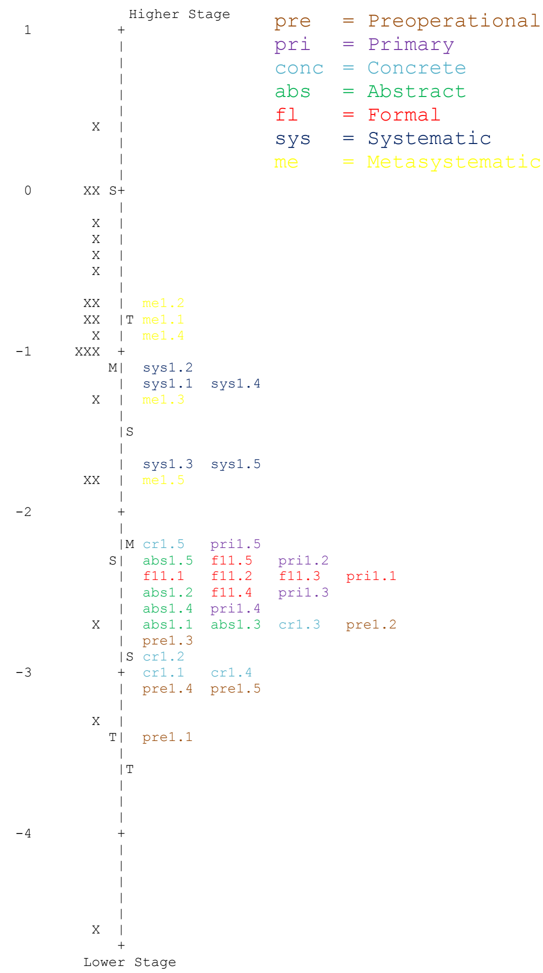


» APPENDIX C1
Empathy rasch map



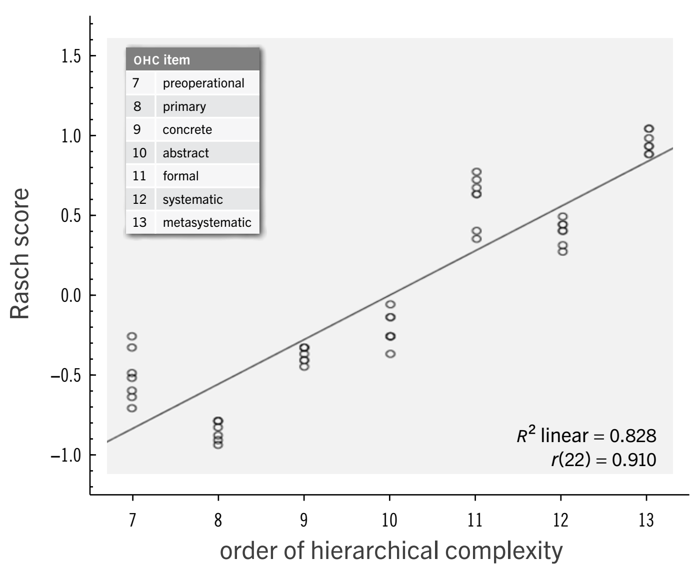
There were 22 participants ranging in age from 18 to 65 (M = 9.00, SD = 2.02) with education varying from high school to graduate school. Each × in the Rasch map represents a participant

» APPENDIX D1
Depression breakup dilemma rasch map

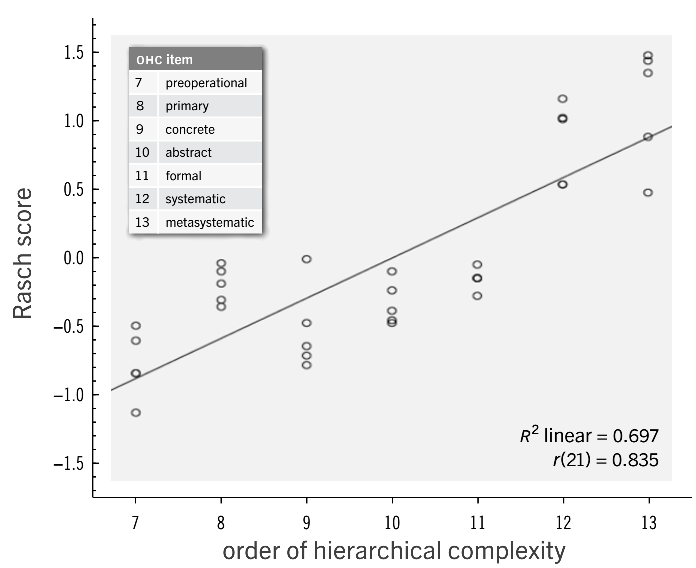


There were 21 participants ranging in age from 18 to 65 (M = 24.07, SD = 9.02) with education varying from elementary school to graduate school. Each × in the Rasch map represents a participant

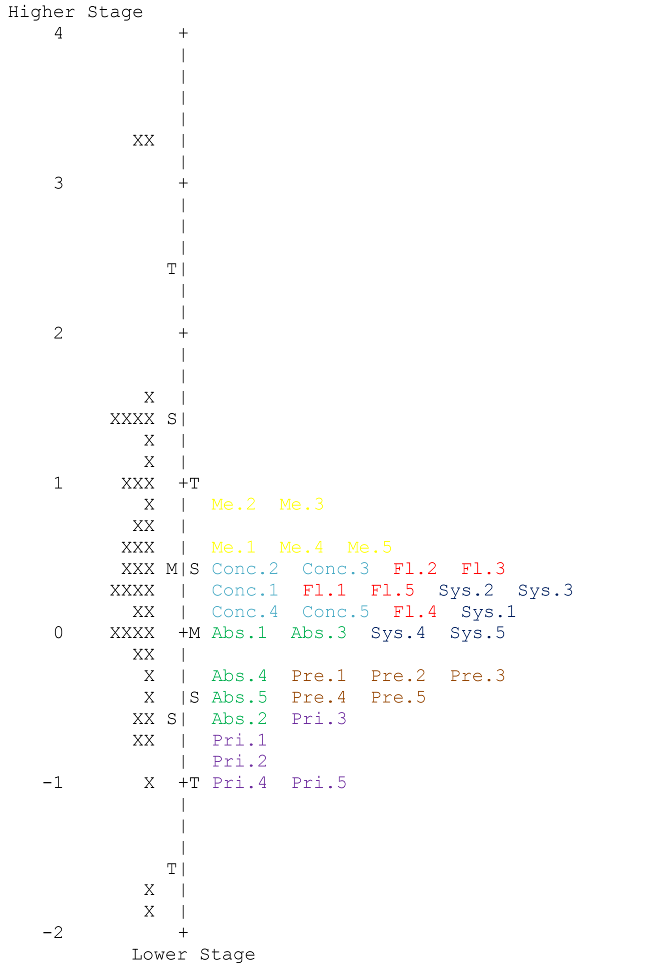
» APPENDIX C2
Empathy results



» APPENDIX D2
Depression breakup dilemma results

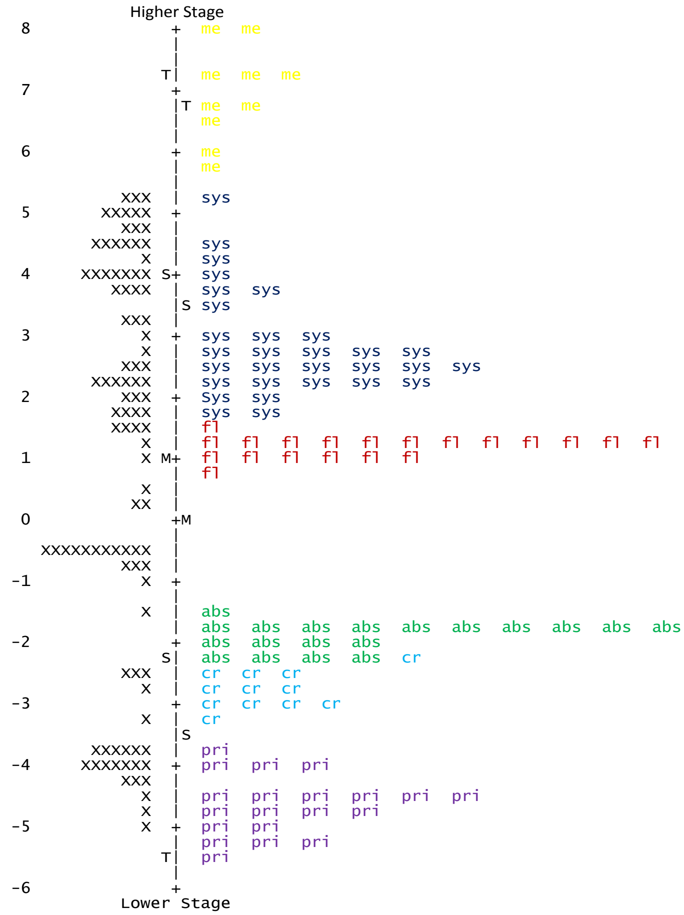


» APPENDIX E1
Caregiver rasch map



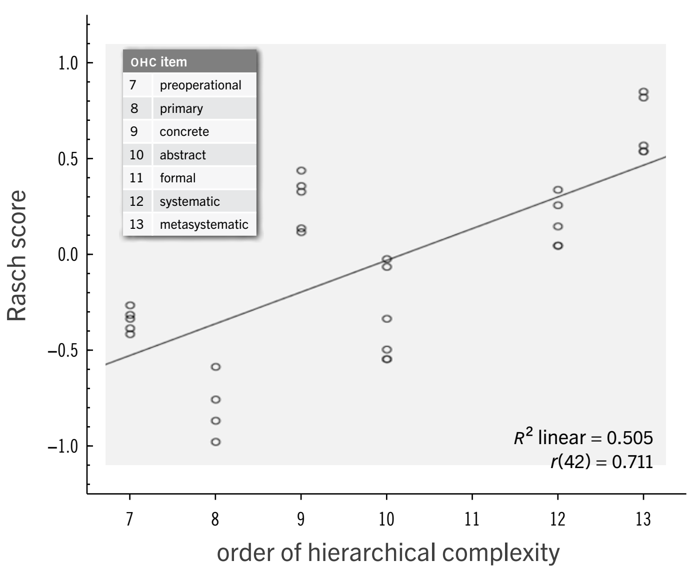
There were 42 participants ranging in age from 19 to 62 (M = 22.90, SD = 8.12) with education varying from elementary school to graduate school. Each × in the Rasch map represents a participant

» APPENDIX F1
Laundry rasch map

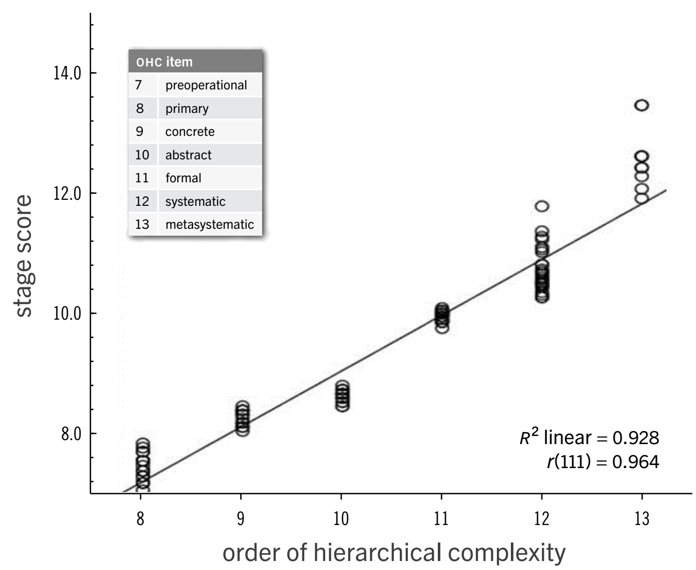


There were 111 participants ranging in age from 12 to 73 (M = 26.02, SD = 10.62) with education varying from elementary school to graduate school. Each × in the Rasch map represents a participant

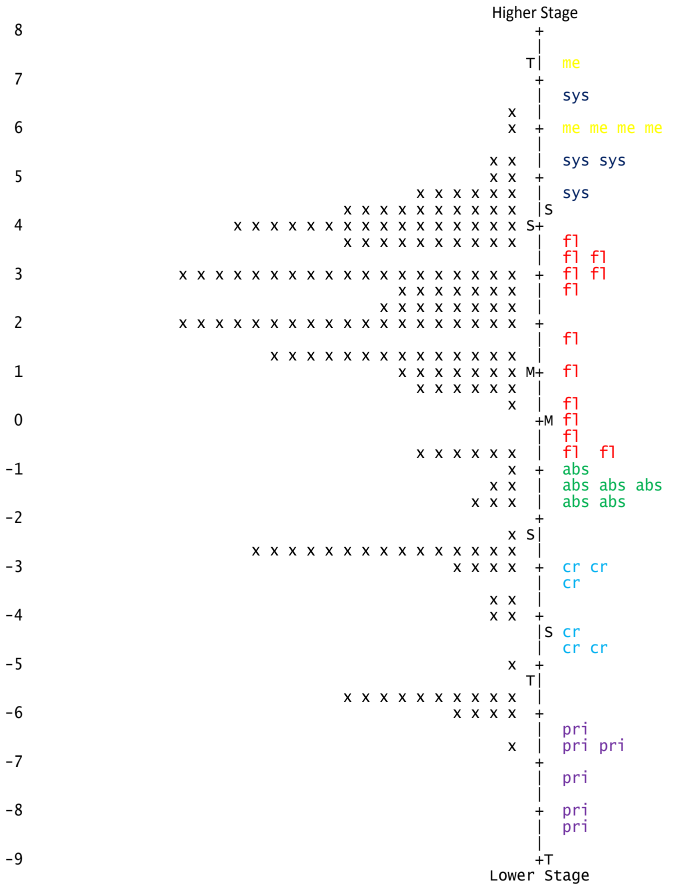
» APPENDIX E2
Caregiver results



» APPENDIX F2
Laundry results

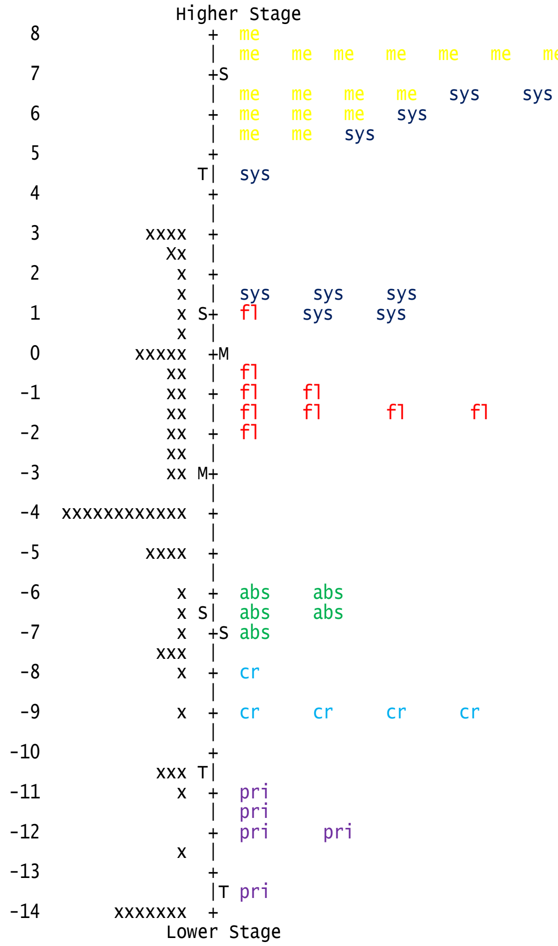


» APPENDIX G1
Algebra rasch map



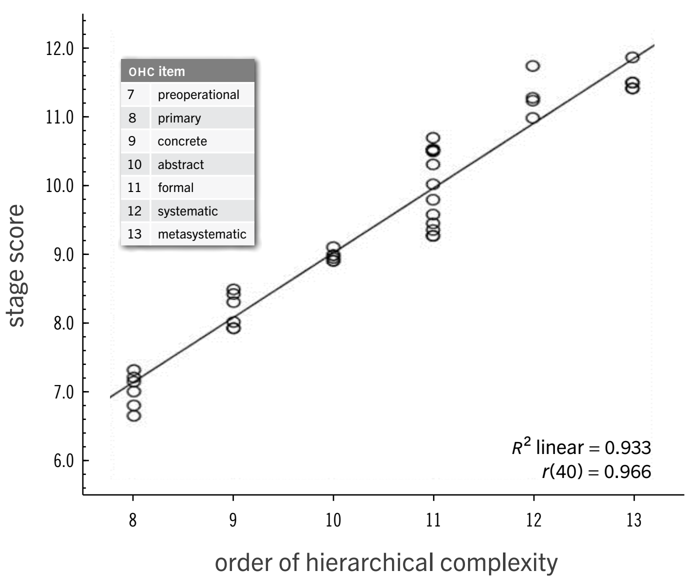
There were 40 participants ranging in age from 14 to 81(M = 24.06, SD = 8.60) with education varying from elementary school to graduate school. Each × in the Rasch map represents a participant.

» APPENDIX H1
Balance beam rasch map

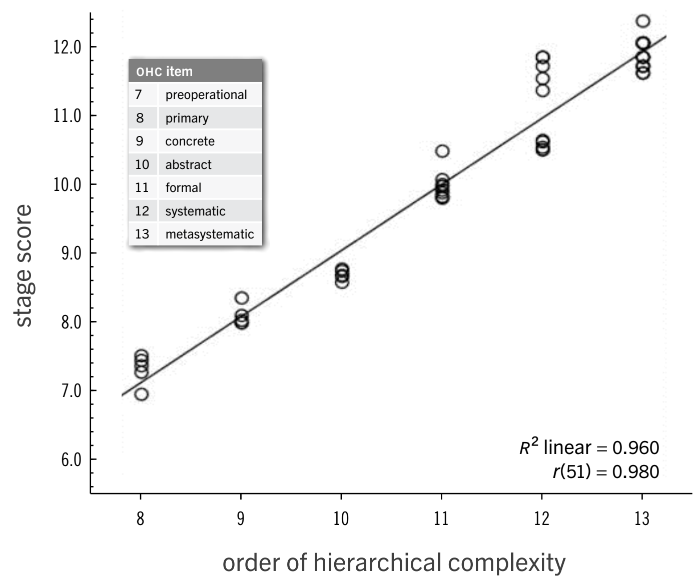


There were 51 participants ranging in age from 12 to 87 (M = 24.11, SD = 9.05) with education varying from elementary school to graduate school. Each × in the Rasch map represents a participant.

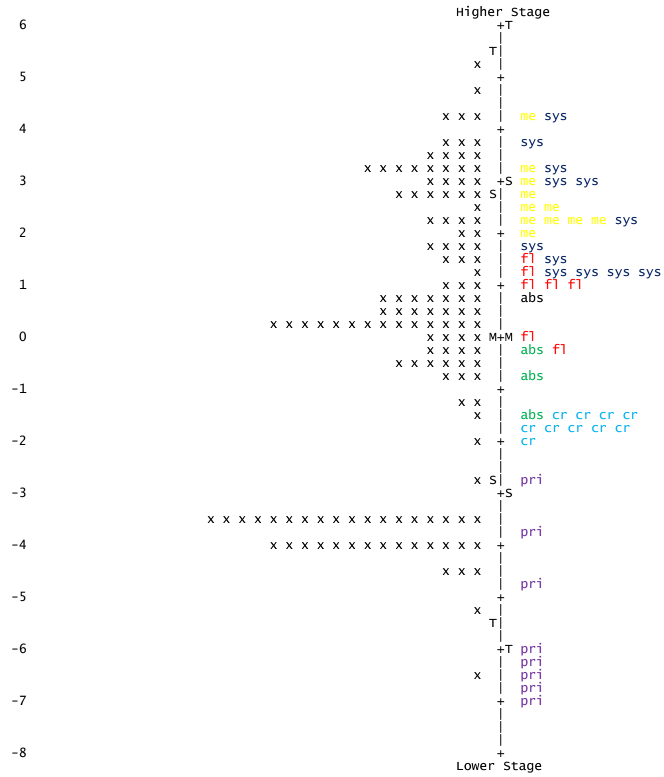
» APPENDIX G2
Algebra results



» APPENDIX H2
Balance beam results

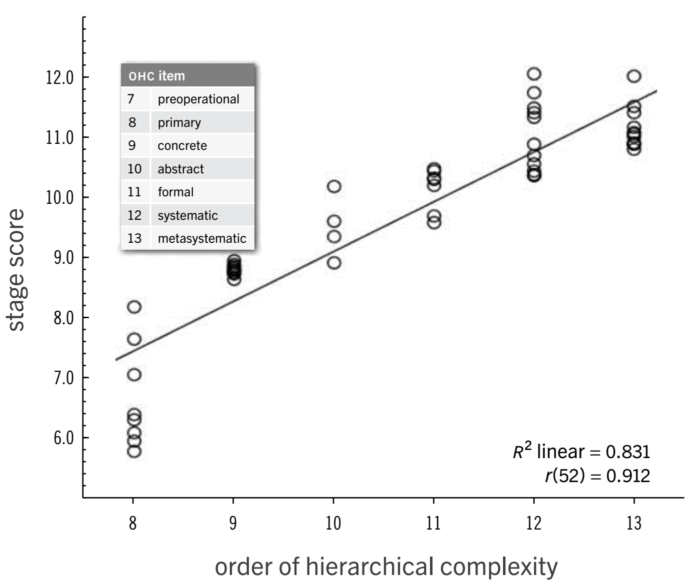


» APPENDIX I1
Infinity rasch map



There were 51 participants ranging in age from 12 to 73 (M = 25.05, SD = 10.51) with education varying from elementary school to graduate school. Each × in the Rasch map represents a participant.

» APPENDIX I2
Infinity results



» APPENDIX J
Social subdomain vignette sequence and questions

Helper person metasystematic vignette example with questions

White speaks with the Person to assess the problem. During the conversation, White offers to provide guidance and assistance seen as most effective in treating this problem. White presents other forms of guidance and assistance as well, and discusses the benefits and risks of each, including doing nothing. White, seeking to understand the Person's needs and concerns, asking and answering many questions. White also sees if the Person's body language matches their statements. White asks if the Person is ready to make a choice based on their previous discussion. Feeling White knows best, the Person accepts the guidance and assistance.

1. Based on the paragraph above, rate each question by selecting a number on the following scales. A rating of 1 means you strongly disagree with Helper. A rating of 6 means you strongly agree with the Helper.

	Extremely poor					Extremely good
	1	2	3	4	5	6
Rate White's method of offering guidance and assistance.	○	○	○	○	○	○
Rate how clearly White expressed their idea.	○	○	○	○	○	○
Rate the degree to which White informed their Person.	○	○	○	○	○	○

2. Rate these two questions on a 1–6 point scale. 1 means you are not at all likely to hire or recommend this Helper, 6 means you are extremely likely to hire or recommend this Helper.

	Not at all likely					Extremely likely
	1	2	3	4	5	6
Rate how likely you would be to accept the guidance and assistance offered by White.	○	○	○	○	○	○
Rate how likely you would recommend White's guidance and assistance.	○	○	○	○	○	○

» APPENDIX J
Logic/mathematics/physical sciences subdomain problem sequence
Balance beam metasystematic problem sequence example

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

ratings	1	x
number of students	$x + 6$	5

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

ratings	500	500y
number of students	$y + 6$	5