DOCUMENT RESUME

ED 428 086		TM 029 486

AUTHOR	Martinez, Michael E.; Simpson, R. Scott
TITLE	Cognitive-Developmental Hierarchies: A Search for Structure Using Item-Level Data.
PUB DATE	1999-02-02
NOTE	29p.
PUB TYPE	Reports - Descriptive (141) -- Tests/Questionnaires (160)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Ability; *Achievement Tests; *Cognitive Development; Cognitive Tests; *Educational Theories; Prediction; *Test Items
IDENTIFIERS	Hierarchical Models; *Proportional Reasoning; *Working Memory

ABSTRACT

Item-level statistics from ability and achievement tests have been underutilized as sources of data for building models of cognitive development. How item data can be used to build a cognitive-developmental map of proportional reasoning is demonstrated. The product of the analysis is a cognitive hierarchy with levels corresponding to categories of cognitive demand established in prior theory. Levels of the constructed hierarchy also correspond to predictions of the working memory demands of the items, and to the measured memory span of examinees. The techniques used in this study could be used more generally to model cognitive-developmental trajectories within other topics and domains. These techniques could be applied to numerous extant assessment databases to identify important conceptual systems in mathematics, science, and other fields, and to construct cognitive-developmental hierarchies within those systems. An appendix contains a test of proportional reasoning. (Contains 2 tables, 2 figures, and 34 references.) (Author/SLD)

Cognitive-Developmental Hierarchies: A Search for Structure Using Item-Level Data

Michael E. Martinez and R. Scott Simpson

University of California, Irvine

Contact: Michael E. Martinez

Department of Education

2001 Berkeley Place

University of California, Irvine

Irvine, CA 92697-5500 USA

Phone: (949) 824-5825

Fax: (949) 824-2965

memartin@uci.edu

Abstract

Item-level statistics from ability and achievement tests have been underutilized as sources of data for building models of cognitive development. In this paper, we demonstrate how item data can be used to build a cognitive-developmental map of proportional reasoning. The product of our analysis is a cognitive hierarchy whose levels correspond to categories of cognitive demand established in prior theory. Levels of the constructed hierarchy also correspond to predictions of the working memory demands of the items, and to the measured memory span of examinees. We propose that the techniques employed in this study could be used more generally to model cognitive-developmental trajectories within other topics and domains. These techniques could be applied to numerous extant assessment databases to identify important conceptual systems in mathematics, science, and other fields, and to construct cognitive-developmental hierarchies within those systems.

Cognitive-Developmental Hierarchies:

A Search for Structure Using Item-Level Data

Data derived from psychometric instruments have long been used to construct theories of human cognition. Developmentalists, for example, have used aggregate test scores to determine a child's mental age, grade-level equivalent, and achievement with respect to an empirically-derived distribution. Information-processing psychologists have studied cognitive components underlying performance on test items and in the process have clarified the nature of such human abilities as induction and visualization (Sternberg, 1985). One class of psychometric data, however, has been neglected as a resource for the development of a cognitive psychology of domain proficiency. This is the item-level and mostly dichotomous information derived from tests of ability, achievement, cognitive style, and personality. Item-level databases are numerous and are often of high quality, consisting of demonstrably reliable information about large and often representative samples of a population. These databases are potentially valuable in helping researchers to understand what is constant and what varies in cognitive development, a quest that has occupied developmental psychologists for at least a century.

The search for regularities in cognitive development has been a tantalizing and often vexing enterprise. Failing to find empirical confirmation of Piaget's stage theory of cognitive development, neo-Piagetian psychologists have largely rejected the proposal that the reasoning processes of a developing child can be accounted for by a central mechanism of logic. Neo-Piagetians have rescued bits and pieces of Piaget's model by maintaining the idea of an orderly development of cognitive proficiency, but in a scaled-back form. The models closest to

Piaget's are those that adhere to a Piagetian generality and parsimony by positing the existence of developmental hierarchies <u>within</u> a small number of powerful thought forms (Case, 1991). According to these models, cognitive development is somewhat independent across thought forms, but within modules orderly development can be traced.

The word <u>module</u>, as used here, refers to a limited knowledge/skill set that has substantive and functional coherence, rather than Fodor's (1985) neurologically-defined sense. It is closer, rather, to Ohlsson's (1993) conceptualization of "abstract schema." One example of cognitive module is the "reasoning pattern," described by Karplus (1981, p. 288) as "an identifiable and reproducible thought process directed at a type of task." Karplus proposed that developmental progressions can be discerned within such thought forms as reasoning about proportionality, causality, and propositional logic. Similarly, Case (1991) has proposed the existence of central conceptual structures modeled as node-link semantic networks, which develop with experience and with efficient use of working memory. Case's prototypical conceptual structure is that of <u>number</u>, but others include reasoning about narratives and n-dimensional space. The models proposed Karplus and Case exemplify theoretical moves toward Piagetian generality and structure, yet stop short of the grand scheme proposed by Piaget himself.

Emergent cognitive-developmental models have much to say to the educator, especially because the newer models grant greater importance to instruction on the developing mind than did Piaget's own theory. If module-based cognitive development has explanatory power, the identity and structure of modules can inform prescriptions for teaching and curriculum design, as well as theories of how children learn. If, for example, there exist a small number of powerful reasoning forms, a lack of proficiency in these forms might constitute quite formidable cognitive impediments, precluding some students from advanced study in certain domains and therefore limiting academic and career attainment. Conversely, instruction targeted toward these cognitive systems and organized around their developmental trajectories could have significant propaedeutic effects, serving to prepare the learner for yet further learning.

The pedagogical utility of key conceptual systems depends significantly on knowledge of their internal structure. Specification of structure should include the essential steps of development within modules, and the relationship of these steps to cognitive processing constraints. Working memory capacity and similar constructs, such as M-power (Pascual-Leone & Baillargeon, 1994), are recognized as important global parameters for cognitive development and as baselines for indexing the cognitive demands of educational tasks (Carpenter, Just, & Shell, 1990; Case, 1991; Halford, 1993), and of cognitive abilities generally (Bereiter & Scardamalia, 1979).

In this study, proportional reasoning is studied as a prototypical conceptual system. Proportional reasoning entails understanding the relationship between at least two variables and applying that relationship to new problems. In problems involving proportionality, a correspondence between ratios is maintained through transformations. Here the operative equation is $A_1/B_1=A_2/B_2$. In science, proportional reasoning is fundamental as exemplified in the core concepts of density (units/space) and time rates (units/time). The pedagogical importance of proportional reasoning is substantial: proportional reasoning has been called "the most ubiquitous mathematical tool of any introductory science course" (Wollman & Lawson, 1978, p. 427). Proportional reasoning ability has also been linked empirically to measured intelligence (Tourniaire & Pulos, 1985).

Although reasoning with part-whole numerical relations has been found in children as young as seven years (Sophian & Wood, 1997), the ability to reason with proportions is by no means universal (Greene, Bsharah, & Bandelow, 1988) at either algorithmic or conceptual levels (Keating & Crane, 1990). Proportional reasoning ability is known to be far less prevalent among disadvantaged students than among those of middle-class backgrounds. Karplus (1981) found, for example, that only between 6% and 16% of urban disadvantaged high school students could reason effectively with proportions, and concluded that "the lack of proportional reasoning by urban high school students would appear to be a serious and tragic obstacle to their education in science and mathematics" (p. 255). Similar observations about the importance of proportional reasoning, and the difficulty of its acquisition by learners, have been made by other investigators (e.g., Resnick & Singer, 1993).

Because of its pedagogical importance, proportional reasoning serves as a module in which to test our methods of analysis. The purpose of this study, then, is to use item characteristics from a test of proportional reasoning to build a hierarchy of proficiency within that domain. We then compare the empirically-derived hierarchy with our prior model derived from a task analysis, as well as the measured memory span of subjects.

Method

Participating students were 46 middle school students in grades 6, 7, and 8. These students were enrolled in an optional 3-week course conducted for off-cycle students in a year-round school. Most participants' families had severely constrained incomes: the program was supported by federal Chapter 1 funds designated for low-income families. Over 90% of participating students were bilingual in English and Spanish, but all participants spoke English with sufficient proficiency that they received instruction in English in their regular classes. In this study, all instruction, except for oral instructions for the tests, was provided in English.

The main object of analysis was a 23-item Test of Proportional Reasoning constructed for this study. Calculators were provided to all students as they solved these problems. The processing demands of problems ranged from simple "protoquantitative compare" operations to the calculation and application of complex ratios (Resnick & Singer, 1993, p. 109). The test items are reproduced and classified by their cognitive demand in the Appendix.

Type I and II items are unidimensional comparison and unidimensional difference items, respectively. Neither represents proportional reasoning proper, but both (comparison and difference) are cognitive precursors that are antecedent to and necessary for true proportional reasoning. For example, pre-proportional reasoning has been conceptualized to involved qualitative differences between states or entities, which are later re-interpreted as metric relations (Resnick & Singer, 1993; Tourniaire & Pulos, 1985). Type I and Type II items represent, respectively, qualitative and quantitative elements of pre-proportional reasoning. Because they involve only two elements, their working memory demands are minimal. It is well understood that working memory is a buffer not only for holding the contents of cognition, but also for carrying out their transformation (Baddeley, 1992). Accordingly, Type I and Type II items are each assigned a working memory load of 3, two for the each of the values compared and one for computation.

By the same logic, simple ratio (Type III) items involve three values and are assigned a working memory load of 4. Type III problems correspond to Level 2 classification presented by Roth and Milkent (1991) and to Piaget Level IIB, in which one of the four ratio terms is equal to one. Simple ratio problems also correspond to those task that are amenable to solution by <u>intuitive</u> as opposed to <u>numerical</u> proportional reasoning (Heller, Ahlgren, Post, Behr, & Lesh, 1989; Moore, Dixon, & Haines, 1991). Complex ratio (Type IV) items involve four values, none of which can be reduced to one, and so must be approached numerically. In one typology (Lamon, 1993) they are referred to as stretchers and shrinkers, because both problem ratios are typically manipulated to obtain a direct comparison. Complex ratio problems are assigned a value of Level 3 and higher in the Roth and Milkent (1991) scheme. Type IV problems are assigned a working memory load of 5, one for each term and one for computation. Similar hierarchical dependencies between proportional reasoning tasks have been cited by other investigators (e.g., Chletsos, De Lisi, Turner, & McGillicuddy-De Lisi, 1989; Lamon, 1993), and these dependencies have been linked to working memory demand (Case, 1991; Roth and Milkent, 1991). Alpha reliability for the Test of Proportional Reasoning was 0.85.

Two tests of memory span, one using a visual digit stimulus and the other using an auditory digit stimulus, were also employed. These tests were reproduced from the ETS Kit of Factor Referenced Cognitive Tests (Ekstrom, French, & Harman, 1976). The memory span tests function as proxies for measures of working memory capacity (Baddeley, 1992). Although working memory and memory span are theoretically distinguishable, both are good predictors of intellectual performance (Halford, 1993), and they are similar theoretically and highly correlated empirically. To insure student comprehension of the tests, spoken instructions were given in both English and Spanish prior to the administration of each test. Students were then allowed to ask questions and receive answers in English or Spanish before beginning each test.

Findings

Our primary objective was to discover developmental trajectories in proportional reasoning on the basis of the psychometric characteristics of test items, in particular item discrimination and identified hierarchical relationships between items. Hierarchical relationships were postulated to represent logical and conceptual dependencies of one item on another (Gagne, 1968). Here we employed the concept of a Guttman scale (Guttman, 1941), which is a psychometric ideal in which items are ordered linearly such that a more difficult item b is never answered correctly unless the response to a simpler item a is also correct (Zwick, 1987). The same contingent relationship holds between items b and c, c and d, up to a theoretical limit of about four items (Bart, 1976).

Hierarchies were constructed by noting Guttman dependencies among the most discriminating items. Guttman orderings were considered ideal when a response to one item is never correct unless a cognitively prior item is answered correctly, implying precedence and developmental dependence. To search for hierarchical relations among items, 2 x 2 frequency tables were constructed for each pair of the 23 items in the Test of Proportional Reasoning. One instance of the frequency tables is shown in Table 1.

Table 1 About Here

In seeking relations of dependency among items in the 2 x 2 crosstabulations, the off-diagonal (lower left and upper right) cells are most telling. In Table 1, the lower left cell shows an entry of zero: None of the 36 students who attempted items 11 and 19 answered item 11 correctly if their answer to item 19 was incorrect. The upper right cell demonstrates the asymmetry of this relationship. Sixteen students (44.4%) gave a correct response to item 19, but not to item 11. These two items were, according to our criteria, a <u>Guttman pair</u>. Our criteria in identifying Guttman pairs specified that one and only one off-diagonal (correct/incorrect) cell would have a value of 0 or 1. A perfect dependency would result in a zero entry in one off-diagonal cell; a frequency of 1, indicating a single reversal of the expected pattern, was tolerated as a "slip" in responding to the easier item. Limited tolerance of slips is a common feature of diagnostic assessment systems (e.g. Bereiter & Scardamalia, 1979; Tatsuoka, 1990) and of Guttman modeling (Hoffman, 1979), and is not incompatible with some theories of cognition which recognize membership in "fuzzy" developmental levels (Moore, Dixon, & Haines, 1991)

A hierarchical structure derived from the crosstabulations might be regarded as epiphenomenal—a necessary consequence of the relative difficulty of items. However, items with comparable differences in difficulty did not always form Guttman pairs. Items 4 and 6, for example, had difficulty values (expressed as proportion correct) of 0.32 and 0.77, but in five cases (out of 39 subjects) subjects answered item 4 (the easier item) correctly, but item 6 incorrectly. The Guttman dependencies we identified, therefore, more clearly reflected ideal discontinuities of distributions in off-diagonal cells than relative difficulty. But even if inter-item dependencies partially reflect underlying difficulty, that leaves open the causes of differential difficulty. Differences in item difficulty are ultimately linked to their information-processing demands. Item difficulty itself has no independent meaning—it merely indexes the item's cognitive demand with respect to a population. Differences in difficulty might, for example, derive from their differential dependence on executive abilities, such as strategic assembly and control of cognitive processes (Snow, 1992), or specific knowledge gained from experience. The item features of interest here are the interdependencies of items, the logical structures underlying performance, and the constraints on cognition imposed by development.

Construction of hierarchies. Two criteria were used to construct hierarchies. One is the criterion of Guttman pairing, described above. The other is each item's discrimination value computed as its correlation with the total test score. According to this assumption, items with high discrimination values are those that share the most variance with composite performance. High discrimination items were assumed to elicit the most broadly applicable cognitive skills for their solutions and to be rich in those components of knowledge, skill, and strategy that Sternberg (1980) called general components. The application of these two criteria, Guttman pairing and high discrimination, resulted in the structure shown in Figure 1.

Figure 1 About Here

In Figure 1, Guttman pairs are indicated by arcs connecting item nodes. The figure indicates Guttman dependencies in all but one of nine pairings between the upper- and lower-tier items. Item/total correlations, which were all relatively high for the items depicted, are printed next to the node representing each item. Printed next to the upper nodes are frequencies of reversals, which are instances in which the upper (more difficult) item was answered correctly and the lower (easier) item incorrectly. Likewise, the frequencies printed beside the lower nodes indicate instances where the lower item is answered correctly, but the upper one incorrectly. If the data conformed perfectly to our model, all integers printed near the upper nodes would be zero and integers near the lower nodes would be moderate in frequency. The pattern shown in Figure 1 approximates this ideal.

The developmental hierarchy was expanded to indicate multiple levels of dependency, which required the inclusion of some less discriminating items. Figure 2 shows the expansion of Figure 1 to incorporate multiple levels. Again, arcs indicate Guttman pairings. The expanded map suggests a pedagogical route and a model for diagnosing the cognitive state of the learner.

Figure 2 About Here

Interpreting the hierarchy. To illustrate the meaning and implications of the cognitive map, consider the items discussed in the previous section, items 19 and 11. Item 19 reads as follows:

> Item 19: Rene's heart beats 70 times per minute. How many times will her heart beat in 3 minutes?

Item 19 is a simple ratio (Type III) problem in which one term is one: 70 beats per one minute. Item 11 is a complex ratio (Type IV) problem:

> Item 11: Robert's recipe for cookie dough calls for five cups of flour and two cups of sugar. Robert decides to use six cups of sugar. How many cups of flour should he use?

In item 11, the ratio which must be scaled up does not contain a unit term; rather, the ratio is 5/2. Even the scaling factor is not given directly as it is in item 19, but must be inferred. Thus, item 19 is logically prior to item 11 in complexity and, presumably, working memory demand. This ordering is reflected in the 2 x 2 tables as well as in Figures 1 and 2.

A general alignment obtains between levels in the hierarchy and our own prior categorization of items according to their cognitive demand. For example, all items in the highest level (IV) correspond to complex ratio problems. Items in the second level (III), with the exception of item 12, are simple ratio problems. Lower in the hierarchy, items 5 and 3 involve unidimensional comparisons (Level I).

Item 4 is an anomaly. It does not fit into the prior categorization scheme and was classified a priori at Level IV. Item 4 reads as follows:

Item 4: Fernando runs 3 miles in 17 minutes; Teddy runs 4 miles in 21 minutes. On average, who ran faster?

The item is complex in structure; our explanation is that its low position is artifactual. Because the response is binary, a student would have a 50% chance of selecting a correct response by random guessing. The guessing factor, we believe, artificially lowered the item's difficulty and resulted in its anomalous and misleading position lower in the hierarchy.

Relationship to memory span. Consistent with some neo-Piagetian models (e.g., Case, 1991; Pascual-Leone, 1970; Roth & Milkent, 1991), we found a correspondence between working memory capacity, approximated by memory span, and levels of the hierarchy. Prior to administering the proportional reasoning items, we assigned to each a hypothetical working memory load by estimating the number of terms and processes that must be held in working memory simultaneously to obtain a correct answer. We later compared the empirical memory span requirements against our task-analytic values. Empirical memory span demand was computed for each item by considering data only for those students who scored a correct response. Thus, for each item shown in Figure 2 we computed the mean and standard deviation of memory span values for those who answered the item correctly. Because half the successful subjects had values lower than the mean memory span, the mean value could not serve as a memory threshold for success. To obtain an estimate of the threshold working memory span requirements to answer the item correctly, two standard deviations were subtracted from the mean of memory span distributions for each item. The resulting values (mean minus two s. d.) were then averaged for items within levels. Table 2 shows that there is a fairly good fit between the predicted and empirical memory span requirements for items within levels of the hierarchy.

Table 1 About Here

Discussion

The methods described in this paper exemplify techniques of data analysis that are useful for identifying and ordering key conceptual systems using the psychometric characteristics of test items. Within domains, such analyses could be used to identify essential knowledge and skill, and for establishing trajectories of developing competency including developmental levels or transitions between levels (Moore, Dixon, & Haines, 1991; Sophian & Wood, 1997). Similar hierachies have been developed in other mathematical subdomains, such as rational number understanding (Kieren, 1993). In the ontogenesis of proportional reasoning, it is clear that a conceptual understanding develops in tandem with improvements in algorithmic proficiency (Keating & Crane, 1990). In our analysis, we relied upon item discrimination values and inter-item dependency relationships to help us structure the more algorithmic and constraint-sensitive side of proportional reasoning. The resulting node-linked structure corresponded to our prior model of domain proficiency and to the memory span of subjects.

The larger significance of this analysis is that readily available psychometric data might be used to illuminate the cognitive structure of other knowledge domains. In this study, item dependencies were identified and developmental hierarchies were constructed on the basis of statistical relations between items. Similar hierarchies might be built using other methods, such as clinical observation of variations in proficiency. Task analyses, also, can yield models of hierarchical relations (e.g., Gagne, 1965, 1968). These different methods of discerning cognitive structure are complementary rather than competitive, because convergent evidence and multiple methods most reliably establish any phenomenon.

One advantage of using the psychometric characteristics of tasks for building cognitive models is that claims about the generalizability of conclusions drawn from such data are likely to be stronger than similar claims based upon clinical samples and techniques. Although some readers may have doubts concerning the validity or value of much current testing practice, the pragmatic importance of achievement and ability tests to life-shaping decisions is beyond question. Also, test items are often much richer in diagnostic potential than is usually appreciated (Bart & Williams-Morris, 1990). Assessment databases, widely available and of demonstrable pragmatic importance, are underutilized for theory building and theory testing by researchers concerned with education. We advocate the use of such data for adding to the knowledge base about learning within subject domains, and have presented one method for doing so.

References

Baddeley, A. (1992). Working memory. Science, 255, 556-559.

Bart, W. M. (1976). Some results of ordering theory for Guttman scaling. Educational and Psychological Measurement, 36, 141-148.

Bart, W. M., & Williams-Morris, R. (1990). A refined item digraph analysis of a proportional reasoning test. Applied Measurement in Education, 3, 143-165.

Bereiter, C., & Scardamalia, M. (1979). Pascual-Leone's M construct as a link between cognitive-developmental and psychometric concepts of intelligence. Intelligence, 3, 41-63.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. Psychological Review, 97, 404-431.

Case, R. (1991). The mind's staircase. Hillsdale, NJ: Lawrence Erlbaum Associates.

Chletsos, P. N., De Lisi, R., Turner, G., & McGillicuddy-De Lisi, A. V. (1989). Cognitive assessment of proportional reasoning strategies. Journal of Research and Development in Education, 22, 18-27.

Ekstrom, R. B. French, J. W., & Harman, H. H. (1976). Kit of factor-referenced cognitive tests. Princeton, NJ: Educational Testing Service.

Fodor, J. A. (1985). Precis of the modularity of mind. The Behavioral and Brain Sciences, 8, 1-42.

Gagne, R. M. (1965). The conditions of learning. New York: Holt, Rinehart and Winston, Inc.

Gagne, R. M. (1968). Learning hierarchies. Educational Psychologist, 6, 1-9.

Greene, A. L., Bsharah, N., & Bandelow, D. (1988). Establishing psychometric equilibrium: The balance puzzle revised. Journal of Adolescent Research, 3, 53-63.

Guttman, L. (1941). The quantification of a class of attributes: A theory and method of scale construction. In P. Horst, et al. (Eds.), The prediction of personal adjustment (pp. 319-348). New York: The Social Science Research Council, Bulletin No. 48.

Halford, G. S. (1993). Children's understanding: The development of mental models. Hillsdale, NJ: Erlbaum.

Heller, P. M., Ahlgren, A., Post, T., Behr, M., & Lesh, R. (1989). Proportional reasoning: The effect of two context variables, rate type, and problem setting. Journal of Research in Science Teaching, 26, 205-220.

Hoffman, R. J. (1979). On testing a Guttman scale for significance. Educational and Psychological Measurement, 39, 297-301.

Karplus, R. (1981). Education and formal thought—a modest proposal. In I. E. Sigel, D. M. Brodzinsky, & R. M. Golinkoff (Eds.), New directions in Piagetian theory and practice. Hillsdale, NJ: Lawrence Erlbaum Associates.

Keating, D. P., & Crane, L. L. (1990). Domain-general and domain-specific processes in proportional reasoning: A commentary on the Merrill-Palmer Quarterly special issue on cognitive development. Merrill-Palmer Quarterly, 36, 411-424.

Kieren, T. E. (1993). Rational and fractional numbers: From quotient fields to recursive understanding. In T. P. Carpenter, E. Fennema, & T. A. Romberg (Eds.), Rational numbers: An integration of research (pp. 49-84). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lamon, S. J. (1993). Ratio and proportion: Connecting content and children's thinking. Journal of Research in Mathematics Education, 24, 41-61.

Moore, C. F., Dixon, J. A., & Haines, B. (1991). Components of understanding in proportional reasoning: A fuzzy set representation of developmental progressions. Child Development, 62, 441-459.

Ohlsson, S. (1993). Abstract schemas. Educational Psychologist, 28, 51-66.

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

Resnick, L. B., & Singer, J. A. (1993). Protoquantitative origins or ratio reasoning. In T. P. Carpenter, E. Fennema, & T. A. Romberg (Eds.), Rational numbers: An integration of research (pp. 107-130). Hillsdale, NJ: Lawrence Erlbaum Associates.

Roth, W.-M., & Milkent, M. M. (1991). Factors in the development of proportional reasoning strategies by concrete operational college students. Journal of Research in Science Teaching, 28, 553-566.

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.

Snow, R. E. (1992). Aptitude theory: Yesterday, today, and tomorrow. Educational Psychologist, 27, 5-32.

Sophian, C., & Wood, A. (1997). Proportional reasoning in young children: The parts and the whole of it. Journal of Educational Psychology, 89, 309-317.

Sternberg, R. J. (1980). Sketch of a componential subtheory of human intelligence. Behavioral and Brain Sciences, 3, 573-614.

Sternberg, R. J. (1985). Human abilities: An information processing approach. New York: W. H. Freeman and Company.

Tatsuoka, K. K. (1990). Toward an integration of item-response theory and cognitive error diagnoses. In N. Frederiksen, R. L. Glaser, A. M. Lesgold, & M. G. Shafto (Eds.), Diagnostic monitoring of skill and knowledge acquisition. Hillsdale, NJ: Lawrence Erlbaum Associates.

Tourniaire, F., & Pulos, S. (1985). Proportional reasoning: A review of the literature. Educational Studies in Mathematics, 16, 181-204.

Wollman, W. T., & Lawson, A. E. (1978). The influence of instruction on proportional reasoning in seventh graders. Journal of Research in Science Teaching, 15, 227-232.

Zwick, R. (1987). Some properties of the correlation matrix of dichotomous Guttman items. Psychometrika, 52, 515-520.

Appendix

Test of Proportional Reasoning

Items Classified by Problem Type and Working Memory Demand

Type I: Unidimensional Comparison (Working Memory Demand=3)

---

1. Tony read 14 books and Sylvester read 12 books. Who read more books?
2. There are 33 houses on Banyan Street and 36 houses on Tresidder Street. Which street has more houses?
5. Bill has three marbles; Antonio has four marbles. Who has more marbles?

Type II: Unidimensional Difference (Working Memory Demand=3)

---

3. The Kings won 4 more games than did the Tigers. If the Tigers won 10 games, how many games did the Kings win?
6. Felicia's car traveled 50 miles and Tom's car traveled 13 miles less than that. How many miles did Tom's car travel?
12. Janet ate three pieces of pizza and Ruben ate six pieces. How many more pieces did Ruben eat than Janet?

Type III: Simple Ratio (Working Memory Demand=4)

---

7. There are ten people in a room and each person has four pockets. How many pockets are in the room?
8. Tina's heart rate is 60 beats per minute. Tina's heart beats 150 times. How much time has gone by?

9. Sylvia rides her bike for one mile and her back wheel goes around 2000 times. How many times will the back wheel go around if she pedals for three miles?

10. In Peter's classroom there is one basketball for every six students. There are five basketballs. How many students are there?

13. Don's drive in puts three pieces of lettuce on each hamburger. If they have 120 pieces of lettuce, how many hamburgers can they make?

14. At the drive-in movies, a car with four adults must pay $16. How much does it cost for one adult?

15. A long distance phone call to El Paso cost $3.25. If the cost per minute was 25 cents, how long was the phone call?

16. When Sandra's family goes to the library, each person chooses six books. If 24 books are checked out, how many people are there in Sandra's family?

19. Rene's heart beats 70 times per minute. How many times will her heart beat in 3 minutes?

Type IV: Complex Ratio (Working Memory Demand=5)

---

4. Fernando runs 3 miles in 17 minutes; Teddy runs 4 miles in 21 minutes. On average, who ran faster?

11. Robert's recipe for cookie dough calls for five cups of flour and two cups of sugar. Robert decides to use six cups of sugar. How many cups of flour should he use?

17. Two tea bags are used to brew three cups of tea. If you want to brew nine cups of tea, how many tea bags should you use?

18. In Mr. Smith's class, there are five boys for every four girls. If the class has 25 boys, how many girls are there?

20. The grocery store is selling three heads of lettuce for $4. If Donella has $20, how may heads of lettuce can she buy?

21. Mr. Rupp bought 21 cookies, and 15 of them had chocolate chips. When he went home, he divided the cookies among his children, making sure that each got the same number of chocolate chip cookies. Timmy got 5 chocolate chip cookies. How many of his cookies do not have chocolate chips?

22. Eli's car traveled 40 miles on 3 gallons of gas. Trina's car went 100 miles on 5 gallons of gas. Whose car has higher gas mileage?

23. Peggy can buy 5 cans of tuna for $3 or she can buy 7 cans for $4. How many cans should she buy if she wants the most for her money?

Table 1

Crosstabulations of Frequencies and Total Percentage for Items 11 and 19, Proportional Reasoning Test (Post-Test)

	Item 19	
Item 11	Incorrect	Correct
Incorrect	10 (27.8%)	16 (44.4%)
Correct	0 (0.0%)	10 (27.8%)

Table 2

Predicted and Actual Memory Span Requirements for Item in the Enhanced Developmental Hierarchy (Figure 2).

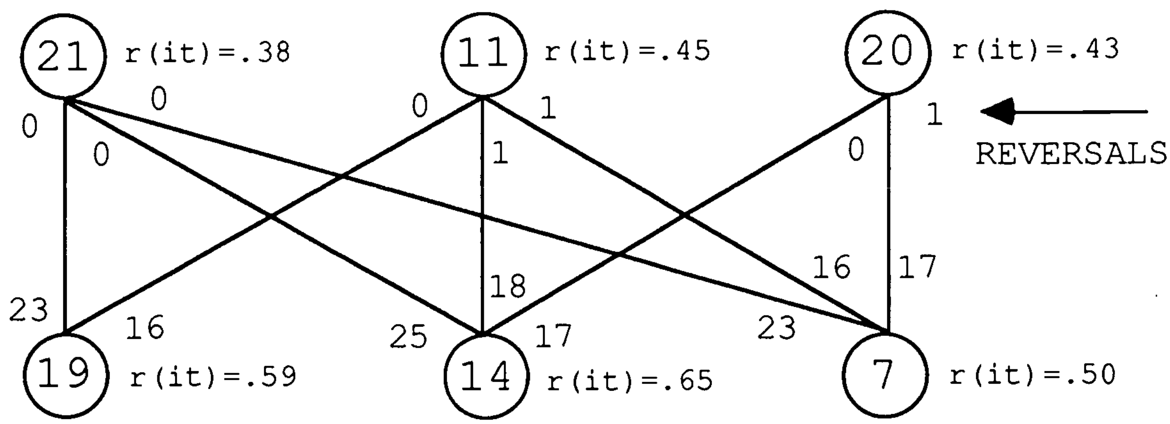
			Memory Span	
Level	Level Name	Items	Predicted	Actual*
IV	Complex Ratios	21,11,20	5	5.12
III	Simple Ratios	16,19,14,8	4	4.11
I	Gross Unidimensional Comparison	5,3	3	3.80

*Actual memory span is the projected minimal memory span requirements. This parameter was obtained by computing the mean auditory memory span minus two standard deviations for subjects who gave a correct response. Mean values were then computed within levels for the items listed.

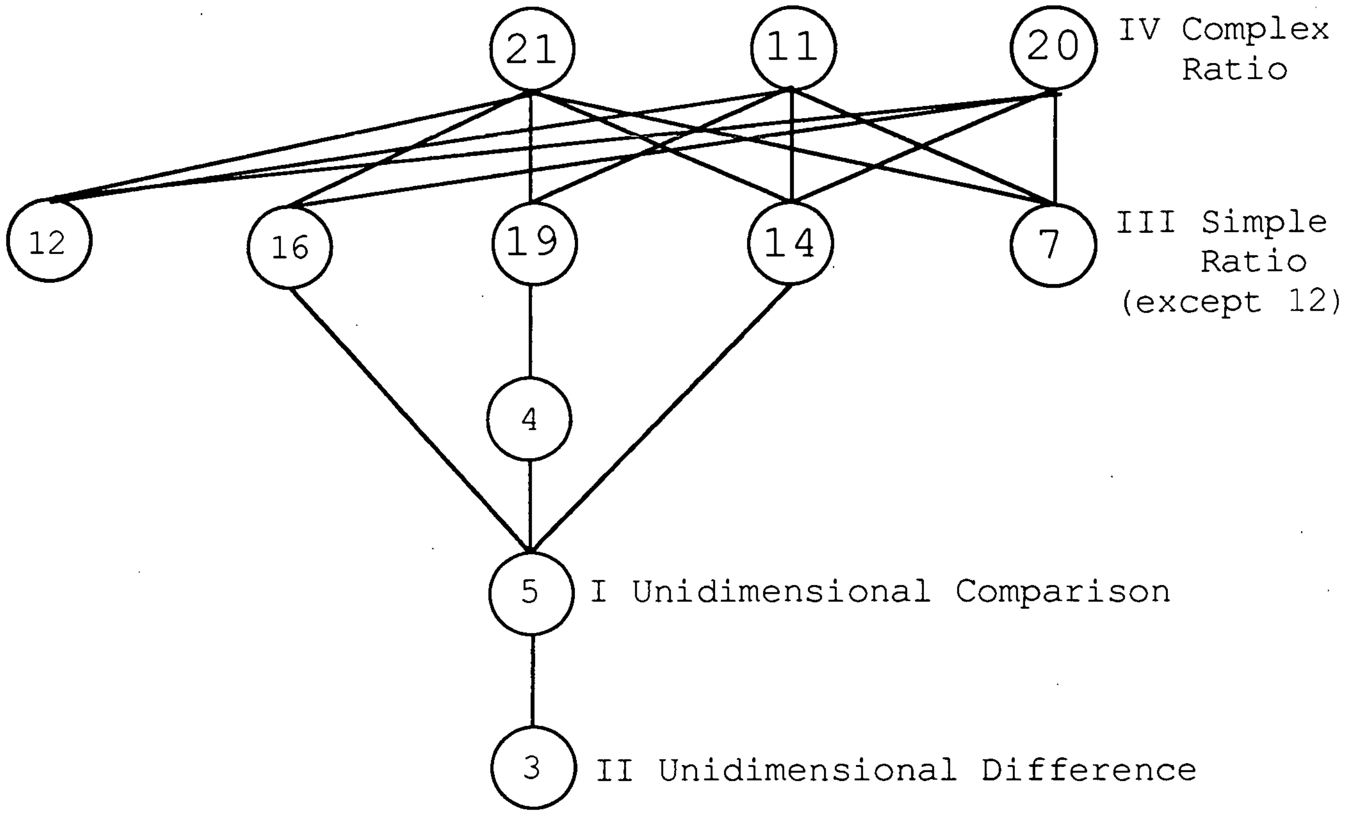
Figure Captions

Figure 1. Hierarchical relationships among key items, with correct/incorrect frequencies noted.

Figure 2. Elaborated cognitive map.



21
r(it)=.38
11
r(it)=.45
20
r(it)=.43
0
0
0
0
1
1
1
0
REVERSALS
16
17
18
23
16
25
17
23
19
r(it)=.59
14
r(it)=.65
7
r(it)=.50


21
11
20
IV Complex
Ratio
12
16
19
14
7
III Simple
Ratio
(except 12)
4
5
I Unidimensional Comparison
3
II Unidimensional Difference