An Ordinal Sequencing Technique for Assessing Multidimensional or Hierarchical Change Models[1]

Walter J. Kuleck and Catharine C. Knight, The Hennepin Group, Inc., 3631 Fairmount Boulevard, Cleveland Heights, OH 44118-4362, and The Department of Educational Foundations and Leadership, The University of Akron, Akron, OH 44325-4208

Abstract. Many scientific disciplines involve the study of growth, development, evolution, or other kinds of multidimensional or hierarchical change processes. The order in which these changes occur can be important to the scientist. Further, understanding these changes may depend on determining not only the order in which they occur, but also the relationships among them. As the broader perspectives now obtaining in science challenge our previous assumptions of linearity and simple sequences, we increasingly require techniques that allow us to view change in a more complex, combining and branching fashion, but with the discipline of statistical rigor.

This paper introduces a statistical technique called "order analysis" to assist both researchers and investigators in any number of different fields in determining 1) both the sequences and the relations among hierarchically changing states or variables based on empirical data, and 2) to compare those sequences and relations to a hypothesized model or theory. Order analysis is a mathematical tool that allows us to identify statistically significant connections between states or variables with assumed ordinality (that is, position in a series or order), and to determine the relative "ordinal distance" between pairs of states or variables so identified. Further, branchings and combinings are identified. This tool allows us, in a relatively simple and straightforward way, to ordinally sequence states and variables, to infer the hierarchical connections to be found between and among states and variables, and to represent the "multidimensional or hierarchical change," "developmental," or "difficulty" distance between identified pairs of states or variables. In this paper we have included an example of the use of order analysis from our literacy research as well as an appendix describing a computer program that implements order analysis.

OHIO J SCI 100 (2):8–12, 2000

INTRODUCTION

In the many types of research that are based on multidimensional or hierarchical change, generally we do not seek data that find our research variables, data or subjects at exactly the same stage or state. Rather, we are seeking to construct models that systematically capture differences in change, growth, development, progress or evolution, either between subjects or within subjects over time. The data we seek reflect the growth of complexity of the subject or process imbedded in one's model of hierarchical change. Of necessity, these data are of differing complexity or hierarchical change from one data point to another, reflecting the underlying model of change. "Hierarchical change" implies that one state follows another in a systematic and orderly fashion, that is, in order, or "ordinally." "Ordinal sequencing" means that one state is expected to come first, another second, yet another third, and so on, as in larva, pupa and adult butterfly. Consequently, the assumptions underlying customary statistical approaches may not be appropriate. For example, the very act of summation, of mean-taking, destroys what to the researcher of change is the essence of the data: the point-to-point or time-to time variation of subject state or performance.

An example of this type of research is found in our study of children's developing literacy skills. Our analytical challenge was to find and refine statistical procedures that document the order, or hierarchy, of the learning that we were assessing. In conducting this work, we came to realize that techniques capable of allowing us to infer ordinal sequences and relationships have utility well beyond psychology and education. Such techniques should be useful in developing multidimensional or hierarchical sequences in other scientific disciplines as well.

METHOD

Order Analysis

We found that a statistical technique called "order analysis" (Bart and Krus 1973; Fischer and others 1993; Krus 1977; Krus and Blackman 1988; Krus and Ceurvorst 1977; Kuleck and others 1990; Tatsuoka 1986) met our identified analytical needs. Order analysis allows the investigator to explore and investigate complex hierarchical change sequences comprising several domains, working together and independently. It permits the sequencing of implicit hierarchies of states or constellations of variables, with multiple developmental pathways between and among states and state areas (or "domains"), including both combining and branching. Consequently, the availability of this relatively new, and still somewhat obscure statistical technique encourages research designs that more fully represent a hierarchical change, or developmental, perspective (Tatsuoka 1986).

[1]Manuscript received 28 December 1998 and in revised form 30 November 1999 (#98-23).

A simple example from education will illustrate the essence of this approach to sequencing ordinal variables, in this case, increasingly difficult knowledge. If two subjects are given the same 100-item biology quiz, and each answers 50 questions correctly, would we give them the same score? Customarily we would do so. This implies that we implicitly (if not always explicitly!) construct such assessments using items of generally equal difficulty. But invariably no two items can be exactly the same. Thus, test items must differ in difficulty, if only very slightly. Strictly speaking we should give our two subjects the same score only if they both get the same 50 items right and 50 items wrong. Of course, we usually accept the hopefully minor inequity caused by our assumption of equal item difficulty and give them the same score regardless.

In our cognitive developmental research, the data we seek reflect the growth of complexity of the skills or learning that we are studying in our research. Since our research is developmental in nature, the assumption of "equal item difficulty" underlying customary statistical approaches simply does not hold (Airasian 1975). Data are necessarily of differing complexity or developmental levels from one data point to another, reflecting an underlying model of growth. Further, the very act of summation, and therefore of mean-taking, destroys what to the researcher of change is the essence of the data: the point-to-point or time-to time variation of subject state or performance (Knight and Kuleck 1987).

Here is the purpose for the order analysis approach. A research model may posit domains (that is, state areas) of a process (a systematic change of state or function) under study, and levels of attainment or progress in each domain. Each level of attainment in a domain we might term a "state." This array of states may at first appear to be a blank slate we seek to fill in with "items" which are tests or indicators of the state. However, we do not know in advance which items are linked or related to which others in the hierarchy of states being represented. We also recognize that we generally cannot devise arrays of states or items whose inter-item developmental interval is uniform on some absolute scale. In short, we need a technique that allows us to preserve point-to-point or time-to-time variation of the state or performance of a subject of study while analyzing inter-item change. Order analysis makes it possible to meet these challenges, as the following example and detailed discussion will show.

RESULTS

Example of the Use of Order Analysis

In our example of our research into children's cognitive development in the realm of developing literacy (Knight and Fischer 1992), we were able to identify several domains, (in this case semantic, phonological, and visualgraphic), and discrete states, or skills, believed to be contained within these domains. Thus, in our research the general hierarchical change matrix shown in Figure 1 was reduced to the specific matrix of six tasks, or items, across three domains reflecting beginning reading skills shown in Figure 2. Figure 3 shows a hypothesized nonlinear developmental sequence for these tasks depicted as a dendrogram, or tree-structured diagram. The top of the diagram is the hypothesized earliest (and by inference, easiest) skill mastered; the bottom is the hypothesized last (and presumably most difficult) skill

	Skill Domains		
Level of Skill	Domain A	Domain B	Domain C
1 (easy)	Skill or Learning A1	B1	C1
2	A2	B2	C2
3	A3	B3	C3
4	A4	B4	...
5	A5	...	...
6	...		
...			
(n) (hard)	An	Bn	Cn

FIGURE 1. Typical skill development matrix. This is an example of a developmental model comprised of three knowledge domains, A, B, and C. At each level of skill, from 1 (easy) to "n" (hard), there is a skill of comparable difficulty in each domain. Both knowledge domain and difficulty may be used to categorize skills. From a developmental perspective, skills A3 and C3, in Domains A & C respectively, are comparable in difficulty and easier than B4, in Domain B.

	Skill Domains			
	Semantic	Visual-Graphic	Phonological	
Skills	Word Definition	Letter Identification	Rhyme Recognition	Easy
		Reading Recognition	Rhyme Production	↓
	Reading Production			Hard

FIGURE 2. Specific hierarchical change matrix for early reading skills. This is a specific example of the general developmental model, using knowledge domains and skills from the authors' research into emergent literacy. Here, "Word Definition," in the Semantic Domain, and "Rhyme Recognition," in the Phonological Domain, are considered to be developmentally equivalent, thus equivalent in "difficulty." These six items could be considered variables in educational or developmental research, or skills in a practitioner environment.

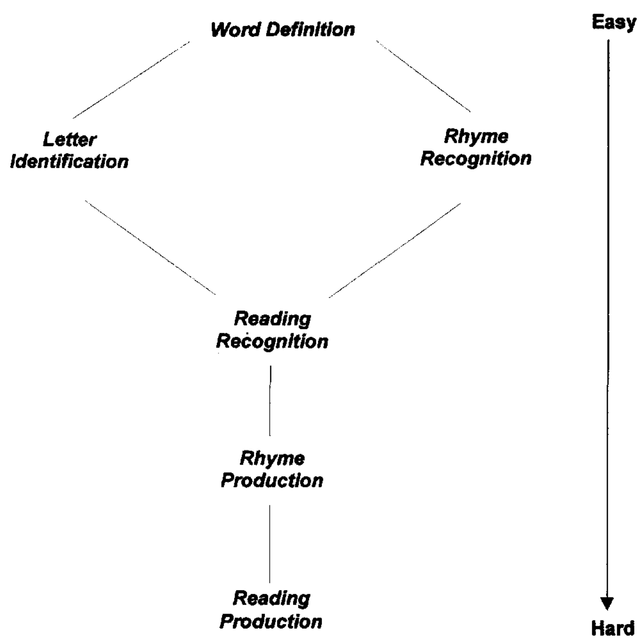

FIGURE 3. Hypothesized ordering of tasks. This "dendrogram" or tree-structure diagram shows the hypothesized orderings and relationships among the six skills or items shown schematically in Fig. 2.

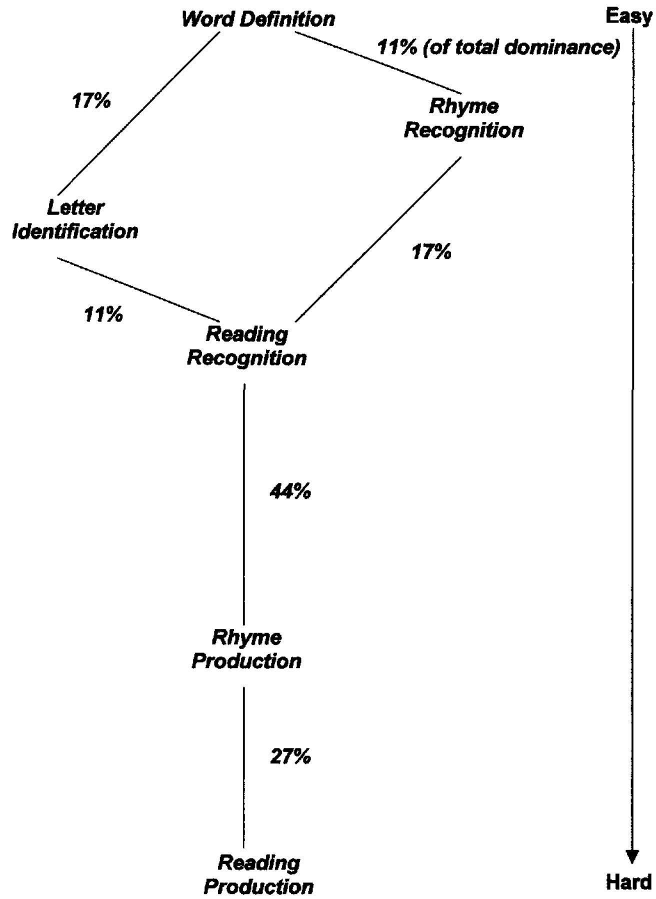

FIGURE 4. Ordering of tasks found with POSI analysis. The results of the POSI analysis in this example verify the hypothesized orderings and relationships shown in Fig. 3. The pairwise dominances are shown as a percentage of the total dominance. The figure is drawn so that the distances between pairs reflect the "developmental distances" or relative difficulty of the items, as inferred from the pairwise dominances. Notice that the shape of the dendrogram is identical to that of the hypothesized ordering in Fig. 3. Thus we infer that the results of order analysis, as shown in Fig. 4, confirm the hypothetical ordering shown in Fig. 3.

mastered. Figure 4 shows the actual ordering derived from the data using order analysis. The order analysis-derived dendrogram (Figure 4) closely matches the hypothesized one (Figure 3); note the branching and combining of tasks is identical in both instances.

Had the actual ordering found been different from that hypothesized, we would have seen a dendrogram in Figure 4 that differed in configuration from that in Figure 3. In this example, no hypotheses were made concerning the magnitude of the "developmental" or "difficulty" distances between skills. The values shown in Figure 4 reflect the percentage of total inter-item dominance (as detailed in the Discussion section following) found from the first skill in the sequence to the last. This manner of dendrogram presentation has become a de facto standard in cognitive developmental research. However, the absolute dominance values may be displayed should they be more meaningful in a particular discipline, or if those values are a part of the researcher's hypotheses.

The results shown in Figure 4, with their inferred complex relationships, could not be obtained or quantified with traditional linear scaling techniques. In this example we were able to glean relationships among skills, or states, and tasks, or items, in developing literacy that had been obscured by traditional techniques.

DISCUSSION

How Order Analysis Works

Order analysis identifies statistically significant connections and the relative hierarchical change distance between each pair of states so connected. First, this technique provides a means to allocate the total variance found in an array of states to each of the state pairs in the array of domains and levels. The total variance found by order analysis is the same as that found by traditional analysis of variance (ANOVA) calculations. The equations, stripped of their matrix algebra trappings, provide the same results as ANOVA. But, instead of calculating variance about the arithmetic mean as in the classic ANOVA approach, order analysis calculates the variance between each pair of states. Since we are dealing with the same differences in both approaches, the reference point for these differences "washes out" as we proceed through the calculations. For more detail on the comparison between order analytical techniques and classical analysis of variance, see Krus and Ceurvorst (1979). The results of order analysis can be portrayed as the proportion of total variance for a given hierarchical change sequence, allocated to each state pairing found.

However, most models of hierarchical or multidimensional change do not routinely assume that every pair of states is a state pairing. For example, a study investigating which patterns of reading skill acquisition foster competent reading in learners may well provide meaningful results. On the other hand, shoe size may seem to be related to one's extent of vocabulary, but this finding may be of little practical or explanatory value. Thus, determining which state pairings (or by inference pairwise hierarchical change sequences) are truly relevant is crucial. Further, how might we statistically test a given pairing; that is, is it a "real" pairing or a chance finding? Order analysis offers some insight into these questions, admitting into sequence only those pairings that are "statistically significant."

Using the example of our developing-literacy research, determining which pairings are significant begins with seeing, for a group of learners, which of each possible pair of skills is found to emerge first (that is, seems easier) for each learner. Referring to Figure 5, if, for example, in a given skill development model, skill G "dominates" (is found to occur prior to, or be a prerequisite for) skill H for five learners (1, 2, 3, 4 & 5) out of fifteen, and H dominates G for none of the learners, we might assume a pair relationship with G linked to H and preceding H in a hierarchical skill development sequence. However, this simple rule, accepting a connection when more learners find one skill in a pair easier and none find the other easier, may be misleading. Thus, we might accept a hierarchical pairing where only one learner, possibly in a very large sample, found one skill to be easier than another in a possible pairing (which we can call a "confirmatory pattern," that is, a pairing consistent with a hypothesis that that skill is easier), so long as there were no learners who found the opposite to be true (a "disconfirmatory" pattern, that is, a pairing contrary to that hypothesis). Likewise, we might reject a hypothesized ordering of a pair of items even, when finding a large number of confirmatory patterns, merely because one disconfirmatory pattern was found. This "rejection" also may be misleading. Simply establishing the presence or absence of disconfirmatory patterns is clearly insufficient. Rather, how many such patterns may we allow, and still infer a hierarchical pairing? Some sort of significance testing is needed to support the finding of any such pairing.

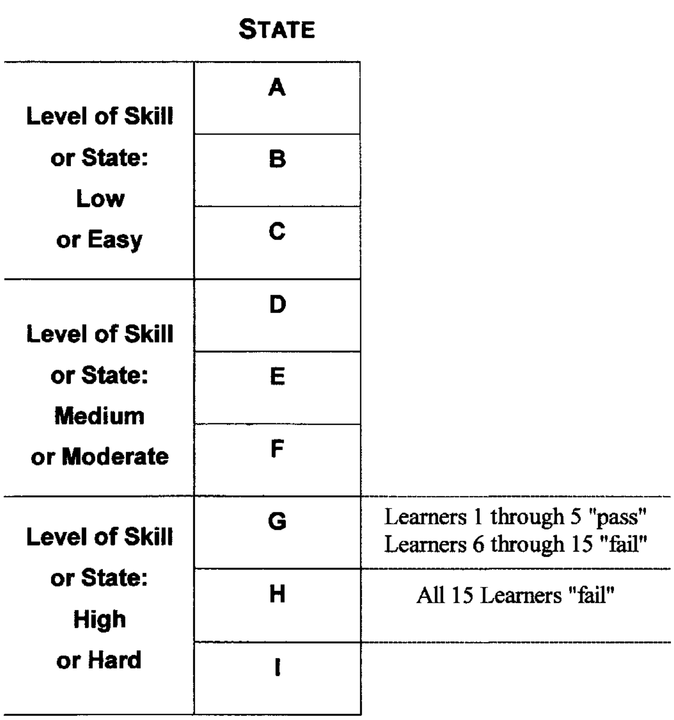

FIGURE 5. Example of dominance: G over H. In this array of nine skills, A through I, A is the "easiest" and I the "hardest." In this example, of the fifteen learners tested, five show mastery ("pass") of skill G, while 10 "fail" skill G. None of the learners "pass" skill H. Thus we may infer that of the skill pair G-H, G "dominates" H, that is, G is "easier" than H and mastery of G may be a prerequisite to mastery of H. This simple example appears unambiguous, as there are no "disconfirmatory" subjects, that is, those that pass H while failing G. But, with real data the pairing will seldom be so clear-cut.

Order analysis, in the implementation known as Partially Ordered Scaling of Items (POSI) (Kuleck and others 1990), uses the probit transformation and the *t*-statistic to help determine statistical significance. POSI is a simple-to-use Windows 95/98 program (Kuleck and Knight 1998) that allows the user to quickly order any array of items into a hierarchical change "tree diagram," or dendrogram. The user simply provides the item responses for the subjects in the study, the desired level of probability and confirmatory frequency, and the range of responses for all the items (typically 0 = fail to 1 = pass). Further details on POSI may be found in the POSI Manual (Kuleck and Knight 1998).

POSI allows us to specify a given probability that the states or items in a given pair are different from one another (which we call level) that we are willing to accept for a given pairing. Note that the "alpha" used here is not the p level used in customary significance testing. This use of alpha refers to the probability that the states or items in a given pair are different from one another. When alpha = 1.00, the inference is that there is no chance the two states are the same, for example, in difficulty. When alpha = .01, only a very small amount of difference is required to consider the two states to be an hierarchical pairing. When the alpha level is allowed to be very small, for example, alpha = 0.01, we will accept nearly all pairings and our dendrogram, or tree diagram of hierarchical changes, will be very broad and complex. If we set alpha = 1.00 we will reject most pairings and the resulting dendrogram will be very simple and restricted. In fact, it may approximate a Guttman scale (Guttman 1944). Krus (personal communication) suggests, and we have found, that a alpha level of alpha = .84 strikes a useful balance between a complex diagram with many pairs that have small differences in difficulty between the items in the pairs, and a sparse diagram caused by requiring very large differences in hierarchy between the states in a pair.

POSI returns an array of both traditional and dominance statistics, as well as a matrix of the pairings found by the analysis and the dominance accounted for by the pairing. Dominance is variance as calculated in a pairwise fashion; the term comes from the concept of one item dominating the other in a pairing. This dominance can be considered a measure of "hierarchical change distance" between the components of each pair. This distance reflects, conceptually, how much change is inferred in going from one item to another paired with it. Hierarchical change distances are additive down a chain of paired items. The sum of the dominances from the most dominant item (the "most dominant" item is defined as that state found to emerge first, the "easiest") to the least dominant (the least dominant item is defined as that state last emerging, the "hardest") will equal, by necessity, the total dominance between the most and least dominant items when the latter is calculated separately.

The results of an order analysis can be plotted as a graphical tree diagram, called a "dendrogram," showing the ordering of states based on the significant hierarchical pairings that were found. In addition, the analysis affords us a measure of "hierarchical change distance," "developmental distance" or "difficulty distance" between states. This distance represents the proportion of the dominance accounted for by a given pairing of states against the total dominance accounted for by all the pairings in that particular hierarchical chain of states.

Figure 4 shows a typical dendrogram with these distances indicated as percentages of the total dominance. Figure 4 also illustrates that perhaps the most exciting inferences that may be drawn from an order analysis are the connections among states that give us, in this example, indications of prerequisites and sequellae of hierarchical skill development. The developmental or difficulty distances place the states in perspective with one another. In Figure 4, it is a small hop from Word Definition to Rhyme Recognition (11% of the total variance of the dendrogram) but a much larger leap from Reading Recognition to Rhyme Production (44% of the total variance of the dendrogram). One pedagogical implication is that more support and instruction would be necessary to help students achieve that larger leap than the smaller hop.

Limitations and Concerns

Successfully sequencing states or variables in a meaningful order depends, first and foremost, on the integrity of the underlying model or theory. This model must have sufficient rigor to lend itself to sensible state or variable definitions and plausible relationships among them. For example, one may sequence items that have no underlying relationship, such as shoe size and hair color, and erroneously believe that this pairing demonstrates a relationship. While no one would seriously propose an ordinal sequence between the two aforementioned variables, more subtly misleading conclusions can be drawn from poorly considered sequencing.

In the case of the development of literary skills used as our example, the underlying model was based on well-accepted theory and research in literacy development. The domains specified are generally accepted as the appropriate state areas for a literacy development model, yielding relatively "pure" states and similar state levels across domains. If the underlying model were to have inappropriate or muddled domains, the sequencing might well be likewise muddled and consequently, of little or no use.

Thus, the researcher proposing to use order analysis or any other ordinal sequencing technique would be well advised to propose a robust model with carefully defined domains and hypothesized processes prior to the actual research. Order sequencing approaches are, after all, essentially descriptive. They can provide a clearer understanding of relationships, but the relationships must be there to describe.

CONCLUSIONS

Developing and testing models of hierarchical or multidimensional change is a complex challenge. We found that the ability practically and statistically to gauge the efficiency of sequencing and presenting states and items was of great help in our literacy research. More generally, we believe that this kind of specification of relations and distances among states can be used to both refine and assess models of hierarchical change. Specifically, in our developing literacy example, we were supported in our expectation that a visual-graphic task (letter identification) and a phonological task (rhyme recognition) develop independently of one another, until an integrative task (recognizing a word from partial cues) is required. Findings such as these can support (or challenge!) assumptions of order and even of ordinality itself.

LITERATURE CITED

Airasian PW. 1975. The linear order assumption and misinterpretation of summary scores. In: Krus DJ, Bart WM, Airasian PW. Ordering theory and methods. Los Angeles: Theta Pr. p 1-7.

Bart WM, Krus DJ. 1973. An ordering-theoretic method to determine hierarchies among items. Edu Psych Measurement 33:291-300.

Fischer KW, Knight CC, Van Parys M. 1993. Analyzing diversity in developmental pathways: Methods and concepts. In: Case R, Edelstein W, editors. The new structuralism in cognitive development: Theory and research on individual pathways. Basel: Karger. p 33-56.

Guttman LA. 1944. A basis for scaling qualitative data. Am Socio Rev 9:139-50.

Fischer KW. 1992. Learning to read words: Individual differences in developmental sequences. J Applied Develop Psych 13:377-404.

Knight CC, Kuleck WJ. 1987. Developmental change: New methods for detection and analysis. Paper given at the annual meeting of the Society for Research in Child Development, April 1987. Baltimore.

Krus DJ. 1977. Order analysis: An inferential model of dimensional analysis and scaling. Educ Psych Measurement 37:587-601.

Krus DJ, Blackman HS. 1988. Test reliability and homogeneity for the perspective of the ordinal test theory. Applied Measurement Educ 1:79-88.

Krus DJ, Ceurvorst RW. 1979. Dominance, information and hierarchical scaling of variance space. Applied Psych Measurement 3:515-27.

Kuleck WJ, Fischer KW, Knight CC. 1990. Partially ordered scaling of items: A program and manual. (Unpublished materials available from the Cognitive Developmental Laboratory, Harvard University, Cambridge, MA.) 28 p.

Kuleck WJ, Knight CC. 1998. Order analysis: Program and manual. Cleveland (OH): Hennepin Group, Inc. 35 p.

Tatsuoka MM. 1986. Graph theory and its applications in educational research: A review and integration. Rev Educ Research 56:291-329.