A Technique for Assessing Effective Instructional Sequences

Catharine C. Knight
The University of Akron

Walter J. Kuleck
The Hennepin Group, Inc.

One of the most complex and challenging responsibilities encountered by educators is developing and structuring effective instruction. As educational researchers we often seek to construct models that systematically capture variations in intellectual growth, skill development, or learning, either between subjects or within subjects over time. As educational practitioners, we endeavor to employ the findings of research and teaching to determine, relate and sequence curriculum and instruction in order to foster the highest quality learning in students. This paper introduces a technique called "order analysis," which we believe can be used to assist both practitioners and researchers in determining those optimal sequences and relations. Order analysis is a mathematical tool that allows us to identify statistically significant connections between items designed to teach or assess skills, and to determine the relative difficulties, called "developmental distance," between skills so paired. Consequently, this tool allows us in a relatively simple and straightforward way to sequence items, to infer the hierarchical connections to be found between and among items, and to represent the "developmental" or "difficulty" distance between item pairs. In this paper, in addition to describing order analysis and its use, we have included an example of its application to our literacy research, as well as an Appendix describing a computer program that implements order analysis.

Assessing Effective Instructional Sequences

In educational research, we often seek to construct models that systematically capture variations in intellectual growth, skill development, or learning, either between learners or within learners over time. In educational practice, among the most fundamental pedagogical tasks for educators are determining, relating and sequencing curriculum, whether basic or advanced, in order to foster the highest quality learning in all students. At first glance this modeling or sequencing might appear to be simple and straightforward. In practice, we have ruefully found that to be the case. Thus, for both researchers and practitioners it would be very helpful to have a means to gauge or assess, statistically and practically, the effectiveness of proposed sequences of instruction toward the development of optimal curricula. Such a technique would be a valuable tool capable of assisting both educational researchers and practitioners in the challenging enterprise of developing optimal learning sequences.

Often our efforts to develop instruction that efficiently sequences concepts and skills for the student are successful. Students learn the material with relative ease and efficiency. But, sometimes the relations among concepts and skills are murky. Learning seems difficult and inefficient. The question emerges: Are we teaching in a manner that is most effective for the learner? Certainly, high quality experience and expertise as educators help us develop workable approaches. Nevertheless, it would be helpful if we could also assess objectively the sequences and relations among skills and concepts to help us best foster optimal learning in our students.

In our research, the authors have faced the task of inferring sequences of concepts and skills learned by students through the study of children's developing literacy skills (Fischer & Knight, 1990; Knight, 1991; Knight & Fischer, 1992). Our 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 inferring sequences and relationships have utility well beyond "academic research." Such techniques should be useful in developing learning sequences in "real life" settings as well.

A Solution: Order analysis

We found that a statistical technique called "order analysis" (Bart & Krus, 1973; Fischer, Knight & Van Parys, 1993; Krus 1977; Krus & Blackman, 1988; Krus & Ceurvorst, 1977; Kuleck, Fischer & Knight, 1990; Tatsuoka, 1986) met our analytical needs. Order analysis allows the investigator to explore complex developmental sequences comprising several skill areas, working together and independently. It permits the sequencing of implicit hierarchies of skills with multiple developmental pathways between and among skills and skill areas, 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).

A simple example will illustrate the essence of this approach to sequencing skill development. If two subjects are given the same 100-item current events quiz, and each answers 50 questions correctly, would we give them the same score? Customarily we would do so, since we implicitly (if not always explicitly!) strive to construct assessments with items of equal difficulty. But, of course, no two items are exactly the same so they 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 generally accept the hopefully minor inequity caused by our assumption of equal item difficulty and give them the same score.

Since our research is developmental in nature, the assumption of "equal item difficulty" underlying customary statistical approaches simply does not hold (Airasian, 1975). Data necessarily differ in complexity or developmental level reflecting the 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 performance (Knight & Kuleck, 1987).

Here is the purpose for the order analysis process. A learning model may posit domains of cognitive process, and levels of attainment in each domain. Each intersection of domain and level we might term a "skill." This array of skills may at first appear to be a blank slate we seek to fill in with "items" which are tests of skill. We do not know in advance which items are linked or related to which others in the student's development, and we cannot devise arrays of items whose "developmental distance," is uniform on some absolute scale. Order analysis makes it possible to meet these challenges.

Example of the Use of Order Analysis

In one example of our research into children's development of developing literacy (Knight & Fischer, 1992), we were able to identify several domains, or relatively delineated skill areas (in this case semantic, phonological and visualgraphic), and discrete 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 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 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. This result, with its 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 and tasks in developing literacy that had been obscured by traditional techniques.

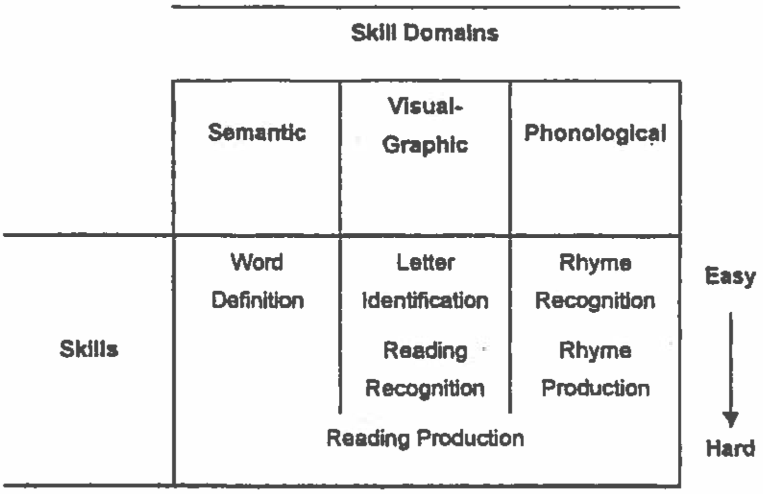
How Order Analysis Works

Order analysis identifies significant connections between skills and the relative difficulties between each pair of skills. This technique provides a means to allocate the total variance found in an array of items to each of the item 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 items. 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 & 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 item pairing found.

However, most models of skill development or learning do not assume that every pair of items is an item pairing. A study of patterns of reading skills that foster competent reading 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 pedagogical value. Determining which item pair (or by inference pairwise hierarchical change sequences) are truly relevant, is crucial. Further, how might we statistically test a given pairing, i.e., is it a "real" pairing or a chance finding? Order analysis offers some insight into these questions admitting into sequence only those pair that are "statistically significant."

Determining which pair are significant begins with seeing, for a group of learners, which of each possible pair of

Figure 1. Typical skill development matrix.



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

Note. 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 skill B4, in Domain B.

items is found to emerge first (i.e., seems easier) for each learner. Referring to Figure 1, 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 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. We might accept a hierarchical pairing where only one learner, possibly in a very large sample, found one skill easier than another (which we can call a "confirmatory pattern), so long as there were no learners who found the opposite to be true (a "disconfirmatory pattern). Likewise, we might reject a pair with a large number of confirmatory patterns when only one disconfirmatory pattern was found. 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.

Order analysis, in the implementation described in the Appendix, uses the probit transformation and the *t*-statistic to help determine statistical significance. In plain language, we may specify a given probability level (which we call alpha level) which 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 skills 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 items are the same, e.g., in difficulty. When alpha=.001, only a very small amount of difference is required to consider the two items an hierarchical pairing. When the alpha level is allowed to be very small, e.g., alpha=.001, 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 a requiring very large differences in difficulty between the items in a pair.

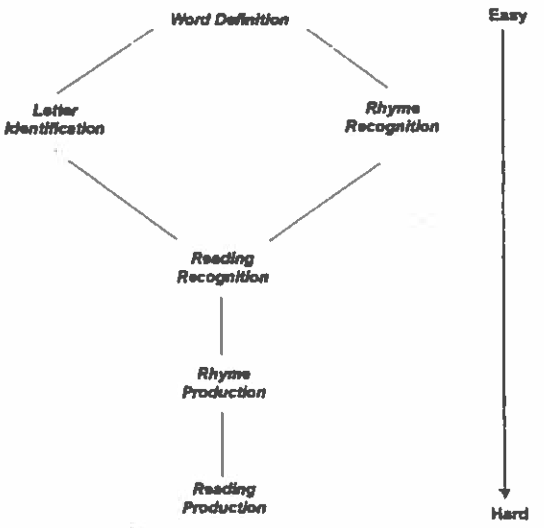
The results of an order analysis can be plotted as a graphical tree diagram, called a "dendrogram," showing the ordering of skills based on the significant hierarchical pairings that were found. In addition, the analysis affords us

Figure 2. Specific Hierarchical change matrix for early reading skills.

	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

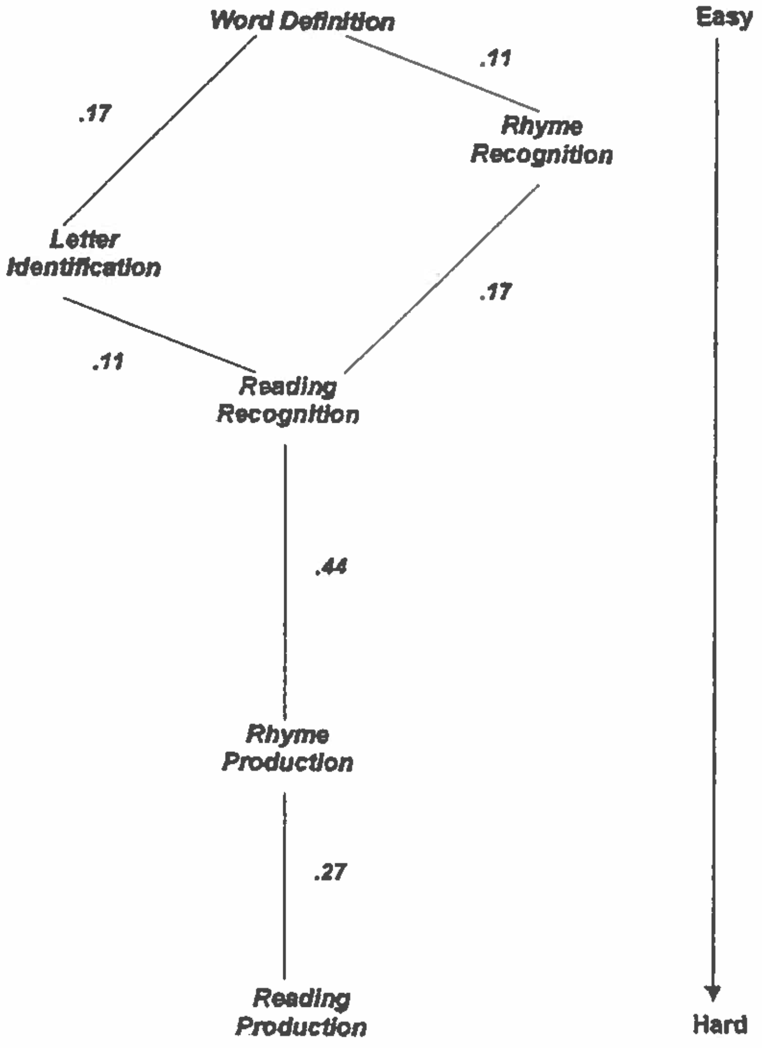
Note. This is a specific example of the general development model shown in Figure 2, using knowledge domains and skills from the authors' research into emergent literacy Knight (1991) and Knight & Fischer (1991). Here, "Word Definition," in the Semantic Domain, and "Rhyme Recognition," in the Phonological 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.



Note. This "dendrogram" or tree-structure diagram shows the hypothesized orderings and relationships among the six skills or items shown schematically in Figure 2.

Figure 4. Ordering of tasks found with POSI analysis.



Note. The results of the POSI analysis in this example verify the hypothesized orderings and relationships shown in Figure 3. The pairwise dominances are shown as a fraction of the total dominance. The figure is drawn so that the distances between pairs reflect the "developmental distance," or relative difficulty of the items, as inferred from the pairwise dominances.

a measure of "developmental distance" or "difficulty distance" between items. This distance represents the proportion of the variance accounted for by a given pairing of skills against the total variance accounted for by all the pairings in that particular hierarchical chain of skills. Figure 4 shows a typical dendrogram with these distances indicated. Figure 4 also illustrates that perhaps the most exciting inferences that may be drawn from an order analysis are the connections among items that give us indications of prerequisites and sequellae of hierarchical skill development. The developmental distances place the skills in perspective with one another. In Figure 4, it's a small hop from Word Definition to Rhyme Recognition (11% of the diagram) but a much larger leap from Reading Recognition to Rhyme Production (44%). One implication is that more support and instruction would be necessary to help students achieve that larger leap than the smaller hop.

Conclusion

Developing and structuring instruction is clearly a complex challenge. We found that the ability to gauge the efficiency of sequencing and presenting skills and tasks was of great help in our literacy research. We believe that this kind of specification of relations and difficulty among concepts and tasks can be used to both expedite and assess instructional development. In our literacy example, we were supported in our expectation that a letter identification and a 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!) current and future instructional development.

References

Airasian, P.W. (1975). The linear order assumption and misinterpretation of summary scores. In Krus, D.J., Bart, W.M., & Airasian, P.W., *Ordering Theory and Methods*. Los Angeles: Theta Press.

Bart, W.M. & Krus, D.J. (1973). An ordering-theoretic method to determine hierarchies among items. *Eduational and Psychological Measurement, 33*, 291-300.

Fischer, K.W., and Knight, C.C. (1990). Cognitive development in real children. In B. Presseisen (Ed.), *Syles of learning and thinking: Interaction in the classroom*. Washington, D.C.: National Education Association.

Fischer, K.W., Knight, C.C. & Van Parys, M. (1993). Analyzing diversity in developmental pathways: Methods and concepts. In Case, R. & Edelstein, W. (Eds.), *The new structuralism in cognitive development: theory and research on individual pathways* (pp. 33-56). Basel: Karger.

Guttman, L.A. (1944). A basis for scaling qualitative data. *American Sociological Review, 9*, 139-150.

Knight, C.C. (1991). Developmental pathways of phonological and visual-graphic prereading skills in young children. Paper given at the Third Annual Convention of the American Psychological Society, Washington, D.C., June, 1991.

Knight, C.C. & Fischer, K.W. (1992). Learning to read words: Individual differences in developmental sequences. *Journal of Applied Developmental Psychology, 13*, 377-404.

Knight, C.C. & Kuleck, W.J. (1987). Developmental change: New methods for detection and analysis. Paper given at the annual meeting of the Society for Research in Child Development, Baltimore, MD, April 1987.

Krus, D.J. (1977). Order analysis: An inferential model of dimensional analysis and scaling. *Educational and Psychological Measurement, 37*, 587-601.

Krus, D.J. & Blackman, H.S. (1988). Test reliability and homogeneity for the perspective of the ordinal test theory. *Applied Measurement in Education, 1*, 79-88.

Krus, D.J. & Ceurvorst, R.W. 1979). Dominance, information and hierarchical scaling of variance space. *Applied Psychological Measurement, 3*, 515-527.

Kuleck, W.J., Fischer, K.W. & Knight, C.C. (1990). *Partially ordered scaling of items: A program and manual.* Cambridge, MA: Harvard University, Cognitive Developmental Laboratory.

Kuleck, W.J. & Knight, C.C. (1998). *Order Analysis: Program and Manual.* Cleveland, OH: The Hennepin Group, Inc.

Tatsuoka, M.M. (1986). Graph theory and its applications in educational research: A review and integration. *Review of Educational Research, 56*, 291-329.

Catherine C. Knight is an Assistant Professor in the College of Education at The University of Akron. Her research and publications have been in the field of cognitive developmental theory.

Walter J. Kuleck is a psychologist in private practice with a major emphasis on analytic and statistical methods.

Appendix

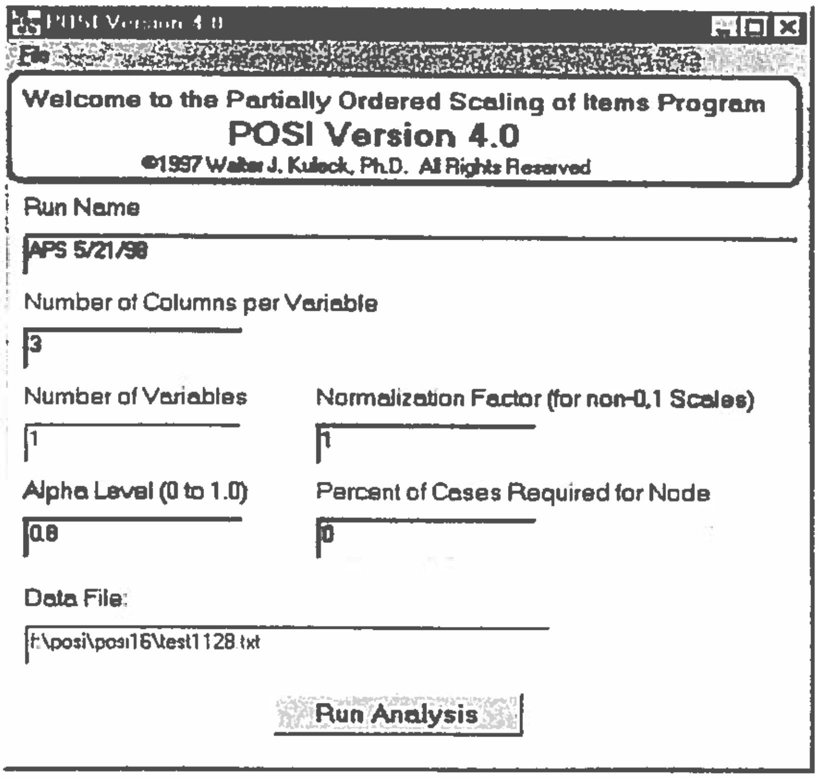
Partially Ordered Scaling of Items Technique (POSI)

The order analytical technique known as Partially Ordered Scaling of Items (POSI) (Kuleck, Fischer & Knight, 1990), is an implementation of order analysis that has been successfully used in a number of hierarchical change research settings. POSI is a simple-to-use Windows 95/98 program (Kuleck & 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). The input screen is shown in Figure A1, with definitions of each field.

The program 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 "developmental distance" between the components of each pair. Developmental distance is, conceptually, how much *harder*one item is from another paired with it. Developmental 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 first mastered or emerging; the "easiest") to the least dominant (the least dominant item is defined as that last mastered or emerging; the "hardest") will equal, by necessity, the total dominance between the most and least dominant items when the latter is calculated separately. Figure 5. Further details on POSI may be found in the POSI Manual (Kuleck & Knight, 1998).

No analysis tool is without its limitations, no matter how useful the tool may be. For example, while POSI may strongly imply the ordering of skills or tasks, actual causality cannot be definitively determined by POSI. Nevertheless, POSI is promising to be a practical tool in understanding developmental pathways particularly in the case when multiple skill domains are being simultaneously assessed, where branching and combining are to be expected.

Figure A1. POSI input screen.



Welcome to the Partially Ordered Scaling of Items Program
POSI Version 4.0
©1997 Walter J. Kuleck, Ph.D. All Rights Reserved

Run Name: APS 5/21/98

Number of Columns per Variable: 3

Number of Variables: 1

Normalization Factor (for non-0,1 Scales): 1

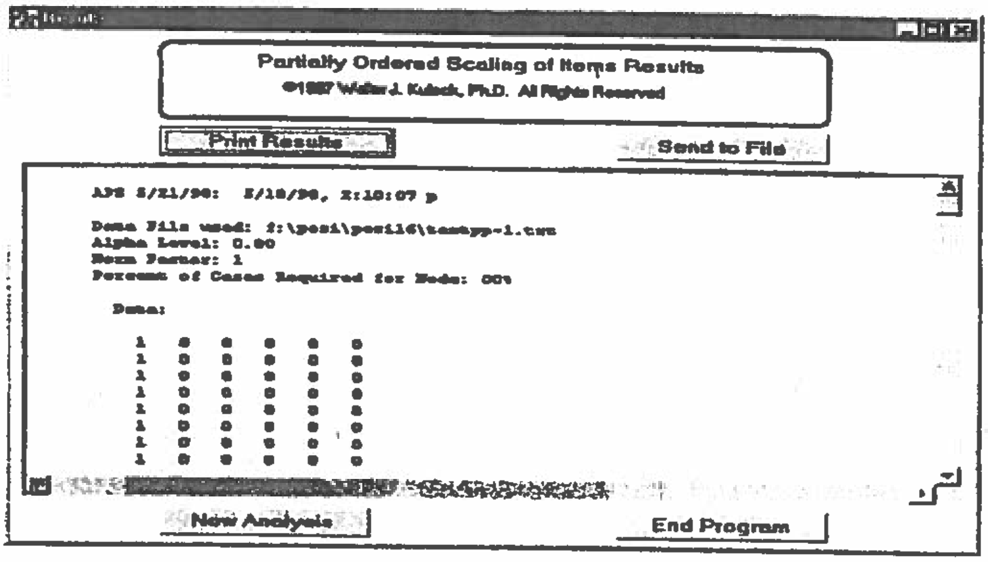
Alpha Level (0 to 1.0): 0.8

Percent of Cases Required for Node: 0

Data File: f:\posi\posi16\test1128.txt

Note. **Run Name**: the name given by the user to identify a particular analysis; **Number of Columns per Variable**: the number of columns in the data file spanned by each variable, or item to be ordered. Data files must be text files with an equal number of columns spanned by each variable; **Number of Variables**: the number of items to be ordered; **Normalization Factor**: for analyses where repeated measures are implicit, i.e., rather than 0 or 1, each variable has multiple responses, e.g., 0 to 4. All variables must have the same range; **Alpha level**: the specified desired level of probability for a pairing to be accepted; **Data file**: the file containing the data to be used for the analysis.

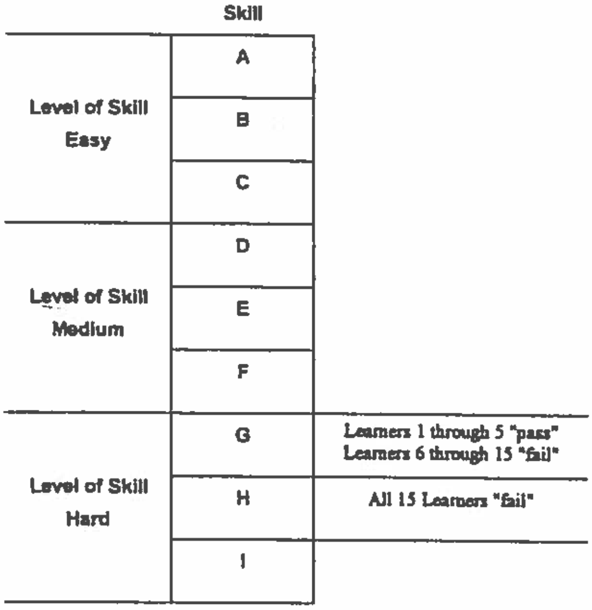
Figure A2. POSI output screen.



```
APS 5/21/98:  5/18/98, 2:10:07 p

Data File used: f:\posi\posi16\testyp-1.txt
Alpha Level: 0.00
Norm Factor: 1
Percent of Cases Required for Node: 00%

 Data:

   1   0   0   0   0   0
   1   0   0   0   0   0
   1   0   0   0   0   0
   1   0   0   0   0   0
   1   0   0   0   0   0
   1   0   0   0   0   0
   1   0   0   0   0   0
   1   0   0   0   0   0
```

Note. Please refer to the POSI instruction manual for a detailed explanation of the program's output.

Figure 5. Example of dominance: G over H.



Note. 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, i.e., 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, i. e. those that pass H while failing G. But, with real data the pairing will seldom be so clear-cut.