

**A RAND NOTE**

**Longitudinal Scalogram Analysis: A Methodology  
and Microcomputer Program for Guttman Scale  
Analysis of Longitudinal Data**

**Ron D. Hays, Phyllis L. Ellickson**

**May 1990**

**RAND**

The research described in this report was supported by the Conrad N. Hilton Foundation and by The RAND Corporation as part of its program of public service.

This Note contains an offprint of RAND research originally published in a journal or book. The text is reproduced here, with permission of the original publisher.

The RAND Publication Series: The Report is the principal publication documenting and transmitting RAND's major research findings and final research results. The RAND Note reports other outputs of sponsored research for general distribution. Publications of The RAND Corporation do not necessarily reflect the opinions or policies of the sponsors of RAND research.

**A RAND NOTE**

**N-3154-CHF/RC**

**Longitudinal Scalogram Analysis: A Methodology  
and Microcomputer Program for Guttman Scale  
Analysis of Longitudinal Data**

**Ron D. Hays, Phyllis L. Ellickson**

**May 1990**

**Supported by the  
Conrad N. Hilton Foundation**

**RAND**



# Longitudinal scalogram analysis: A methodology and microcomputer program for Guttman scale analysis of longitudinal data

RON D. HAYS and PHYLLIS L. ELLICKSON  
*The RAND Corporation, Santa Monica, California*

Guttman scalogram analysis has been limited to cross-sectional analysis. Longitudinal scalogram analysis (LSA), a direct extension of cross-sectional scalogram analysis to longitudinal data, is proposed as an alternative methodology. The benefits of LSA relative to cross-sectional methods of analysis are discussed.

Unitary growth characterizes a variety of different developmental processes including intellectual development (e.g., see Bayley, 1955), drug use involvement (e.g., see Kandel, 1975), moral development (Walker, de Vries, & Bichard, 1984), and functional health (Stewart, Ware, & Brook, 1981). The common feature of these different domains is a deterministic, cumulative sequence of development. Cross-sectional Guttman scale analysis has been employed as the bread-and-butter method for evaluating these processes (Guttman, 1944).

If dichotomous items are consistent with a Guttman scale, then they are ordered along a single dimension (in terms of the prevalence of their endorsement) and the pattern of responses to individual items is determined by the sum of the endorsed items. Therefore, respondents with the same total score should have identical responses to the items in a Guttman scale. Although the number of possible response patterns is two raised to a power equal to the number of items, the number of response patterns consistent with a Guttman scale equals the number of items plus one (Dotson & Summers, 1970; Schwartz, 1986).

Table 1 presents the item response patterns expected for three items forming a Guttman scale of measurement; *magnitude*, *equal interval*, and *absolute zero*. Eight response patterns are possible, but only the four shown in Table 1 are consistent with a Guttman scale. Knowing that a scale has an absolute zero point allows for the inference that it has equal intervals and that it has the property of magnitude or "moreness" (see Kaplan & Saccuzzo, 1989). Similarly, knowing that a scale has equal intervals leads to the prediction that it possesses the property of magnitude. In contrast, knowing that a scale has magnitude does not allow one to infer whether or not it has equal intervals or an absolute zero point.

Preparation of this article was supported by a grant from the Conrad N. Hilton Foundation and by the RAND Corporation as part of its program of public service. The opinions expressed are those of the authors and do not necessarily reflect the views of the RAND Corporation. Requests for reprints may be sent to Ron D. Hays or Phyllis L. Ellickson, RAND Corporation, 1700 Main Street, Santa Monica, CA 90406-2138.

Table 1  
 Example of Pattern of Responses to Three Items Fitting Perfectly a Cross-Sectional Guttman Scale

Type of Person	Magnitude?	Equal Interval?	Absolute Zero?	Total Score
Nominal	NO	NO	NO	0
Ordinal	YES	NO	NO	1
Interval	YES	YES	NO	2
Ratio	YES	YES	YES	3

The scalability of responses is determined by comparing observed patterns of data with the patterns predicted for a Guttman scale, examining the degree to which observed response patterns deviate from expected response patterns. The *coefficient of reproducibility* (CR) for Guttman scales is defined as the proportion of error (i.e., proportion of differences between observed and expected responses), subtracted from unity. A CR value of 0.90 or higher is considered acceptable. In addition, an index of reproducibility is typically computed by determining how well item modes reproduce the observed response patterns. (Errors are counted as differences between each observed item response for an individual and the modal response for that item across all respondents.) This index, the *minimum marginal reproducibility* (MR), is used to calculate the *coefficient of scalability* (CS), defined as  $(CR - MR)/(1 - MR)$ . A CS of 0.60 has been recommended as a minimum standard for acceptability (Menzel, 1953).

## Longitudinal Scalogram Analysis

Traditional Guttman scalogram analysis is limited to evaluating item order cross-sectionally. An extension of traditional Guttman scaling that incorporates the element of time, the Longitudinal Guttman Simplex (LGS), has been developed by Collins, Cliff, and Dent (1988). They defined a consistency index for longitudinal analysis, CL, as follows:

$$CL = \frac{\text{Observed} - \text{Expected}}{1 - \text{Expected}}$$

Reprinted from *Behavior Research Methods, Instruments, & Computers*, N. John Castellan, Jr. (ed.), Vol. 22, No. 2, pp. 162-166, © 1990 by The Psychonomic Society. Reprinted by permission.

**Table 2**  
Four Types of Item-Time Relations in the Longitudinal Guttman Simplex

	Item Pair							
	A		B		C		D	
	1	2	1	2	1	2	1	2
Time 1	NO	NO	YES	NO	NO	YES	NO	YES
Time 2	YES	YES	YES	NO	YES	YES	YES	NO

Note—Pair A represents a redundant time relation, Pair B a redundant item relation, Pair C a unique relation, and Pair D a contradictory relation.

**Table 3**  
Example of Pattern of Responses to Three Items Fitting Perfectly a Three-Wave Longitudinal Guttman Scale

A1	B1	C1	A2	B2	C2	A3	B3	C3	Total Score
NO	NO	NO	NO	NO	NO	NO	NO	NO	0
NO	NO	NO	NO	NO	NO	YES	NO	NO	1
NO	NO	NO	YES	NO	NO	YES	NO	NO	2
NO	NO	NO	NO	NO	NO	YES	YES	NO	2
NO	NO	NO	YES	NO	NO	YES	YES	NO	3
NO	NO	NO	NO	NO	NO	YES	YES	YES	3
YES	NO	NO	YES	NO	NO	YES	NO	NO	3
NO	NO	NO	YES	YES	NO	YES	YES	NO	4
NO	NO	NO	YES	NO	NO	YES	YES	YES	4
YES	NO	NO	YES	NO	NO	YES	YES	NO	4
NO	NO	NO	YES	YES	NO	YES	YES	YES	5
YES	NO	NO	YES	YES	NO	YES	YES	NO	5
YES	NO	NO	YES	NO	NO	YES	YES	YES	5
NO	NO	NO	YES	YES	YES	YES	YES	YES	6
YES	YES	NO	YES	YES	NO	YES	YES	NO	6
YES	NO	NO	YES	YES	NO	YES	YES	YES	6
YES	NO	NO	YES	YES	YES	YES	YES	YES	7
YES	YES	NO	YES	YES	NO	YES	YES	YES	7
YES	YES	NO	YES	YES	YES	YES	YES	YES	8
YES	YES	YES	YES	YES	YES	YES	YES	YES	9

Note—A = magnitude, B = equal interval, C = absolute zero.

“Observed” is the weighted proportion of consistent relations between items across time that are congruent with the a priori specified ordering of items and times. “Expected” is the weighted proportion of consistent relations expected by chance.

Four kinds of relations of items and times are examined: redundant (time and item), unique, and contradictory. *Redundant time relations* are those in which answers given to a pair of items provide redundant information about

two time points (i.e., at one time point, both items are failed and at the other time point both are passed). *Redundant item relations* are those in which the answers to a pair of items match at two time points. *Unique relations* are those in which responses to only one item in a pair changes over time. *Contradictory relations* provide conflicting information about the relative ordering of both items and times.

Table 2 presents each of the four types of relations possible between pairs of items and pairs of times. Pair A represents a redundant time relation. Horizontal lines connect responses that concur. Pair B depicts a redundant item relation. Vertical lines connect responses that agree. Pair C is an example of a unique relation. One horizontal and one vertical line link responses that agree. Pair D illustrates a contradictory relation. Diagonal lines connect responses that concur.

CL is a quality index that ranges from negative infinity to positive one (see Cliff, 1979). The weighting scheme used to compute CL was empirically derived on the basis of the ability to distinguish random from nonrandom data and to distinguish among data known to differ in consistency (Collins et al., 1988). Unique relations are weighted four times as much as redundant and contradictory relations. The total number of weighted consistent relations is computed as the sum of redundant and four times the number of unique relations that are congruent with the a priori item-times order. The proportion of consistent relations is equal to the total number of consistent relations divided by the total number of weighted relations (cf. Collins et al., 1988). Rules of thumb for the CL index have been suggested, but precise interpretation guidelines have not been developed (Collins & Cliff, in press; Collins & Dent, 1986).

We propose an alternative longitudinal analysis methodology, longitudinal scalogram analysis (LSA), which is a direct extension of cross-sectional Guttman scalogram analysis. Table 3 presents patterns of responding for three items measured at three time points. As is illustrated in Table 3, only one pattern of responses is longitudinally consistent with a total score of 0, 1, 8, or 9 “yes” answers. However, there are two different response patterns consistent with 2 or 7 affirmative answers and three different response patterns consistent with a total of 3, 4, 5, or 6 affirmative answers. For example, a total score of 2 may be obtained for a scale having the property of *magnitude* at Time 2 and Time 3 or by a scale having *magni-*

**Table 4**  
Comparing Example Pattern to Patterns Consistent with a Longitudinal Guttman Scale: Three Items, Three Waves, and a Total Score of 5

Time 1	Items			Difference Between Patterns	
	Time 2	Time 3			
	1 2 3	1 2 3	1 2 3		
0 0 1	1 1 1	1 0 0			Example Pattern
0 0 0	1 1 0	1 1 1	4		Longitudinally Consistent Pattern 1
1 0 0	1 1 0	1 1 0	4		Longitudinally Consistent Pattern 2
1 0 0	1 0 0	1 1 1	6		Longitudinally Consistent Pattern 3

Note—0 = not passed, 1 = passed.

Table 5  
Example Input to LSA.EXE

```

NCASES
6
WAVES
2
HOWREAD
(611)
VARIABLES
6
  LOW2  1
  MED2  2
  HIGH2 3
  LOW1  4
  MED1  5
  HIGH1 6
DATA
000000
100100
110110
111111
101101
011011
END

```

Table 6  
Primary Output of LSA.EXE

ITEM	PROPORTION PASSING
Wave = 2	
1	0.67
2	0.50
3	0.50
Wave = 1	
1	0.67
2	0.50
3	0.50
NUMBER OF SUBJECTS	= 6
LONGITUDINAL SCALOGRAM ANALYSIS	
COEFFICIENT OF REPRODUCIBILITY (MAX)	= 0.889
COEFFICIENT OF SCALABILITY (MAX)	= 0.750
COEFFICIENT OF REPRODUCIBILITY	= 0.833
MINIMUM MARGINAL REPRODUCIBILITY	= 0.556
PERCENT IMPROVEMENT	= 0.278
COEFFICIENT OF SCALABILITY	= 0.625
CROSS-SECTIONAL SCALOGRAM ANALYSIS	
COEFFICIENT OF REPRODUCIBILITY WAVE 2	= 0.778
MINIMUM MARGINAL REPRODUCIBILITY	= 0.556
PERCENT IMPROVEMENT	= 0.222
COEFFICIENT OF SCALABILITY	= 0.500
COEFFICIENT OF REPRODUCIBILITY WAVE 1	= 0.778
MINIMUM MARGINAL REPRODUCIBILITY	= 0.556
PERCENT IMPROVEMENT	= 0.222
COEFFICIENT OF SCALABILITY	= 0.500

tude and equal interval properties at Time 3 (assuming scales can change over time). Because of the multiple response patterns consistent with a longitudinal Guttman scale, the calculation of reproducibility and scalability is not as straightforward for longitudinal as it is for cross-sectional data.

With longitudinal data, the expected pattern against which observed scores are compared cannot be determined solely on the basis of the total score across items. However, identification of all longitudinal patterns that are consistent with the Guttman model and yield the total score observed for each individual can be used to select the pattern (i.e., expected pattern) that is *minimally different* from observed scores. Table 4 provides an example of selecting the expected pattern for a total score of 5 and observed score pattern of 001 111 100 for three items measured at three time points. The minimum difference between the observed pattern and the three patterns consistent with a longitudinal Guttman scale and yielding the same total score is 4. This difference is observed for two of the three patterns; thus, either of these patterns can serve as the expected pattern (i.e., they are equivalent for the purpose of computing scalogram errors).<sup>1</sup>

Once the expected pattern has been determined, longitudinal coefficients of reproducibility (LCR) and scalability (LCS) can be computed as in cross-sectional Guttman scalogram analysis. Subtracting the proportion of errors from unity yields LCR. LCS is defined as the difference between the reproducibility of items from their modes (LMR) and LCR, divided by LMR subtracted from unity:  $LCS = (LCR - LMR)/(1 - LMR)$ .

A microcomputer program, LSA.EXE, has been developed to compute longitudinal coefficients of reproducibility and scalability. LSA.EXE 1.0 is a compiled BASIC program that runs under MS-DOS Version 2.0 or later on IBM PC or compatible (e.g., COMPAQ) microcomputers. LSA.EXE 1.0 is limited to four waves (time points) of data.

*How to Use LSA.EXE.* A user-supplied ASCII file of 0s and 1s is required input to LSA.EXE 1.0. LSA.EXE 1.0 outputs the proportion of the sample passing each item, the number of respondents in the analysis, a frequency distribution of the number of scaling errors, and the longitudinal coefficients of reproducibility and scalability, LCR and LCS. Cross-sectional coefficients of reproducibility and scalability are provided for each wave of data. In addition, the universe of response patterns perfectly consistent with a longitudinal Guttman scale for the given number of items and waves is printed, sorted by the number of endorsed items. Table 5 provides an example of the input setup needed to run LSA.EXE. Tables 6-8 give examples of the program's output.

### Conclusions

Despite the interest in charting development over time, in most studies, cross-sectional Guttman scalogram analysis has been used to draw inferences about transitions. Cross-sectional Guttman scale analysis does not permit

Table 7  
Supplementary Output of LSA.EXE

FREQUENCY OF SCALING ERRORS	
0 **** ( 4)	
2 * ( 1)	
4 * ( 1)	
NCASES=	6
NWAVES=	2
HOWREAD=	(611)
VARIABLES=	3
LOW2	1
MED2	2
HIGH2	3
LOW1	4
MED1	5
HIGH1	6

Table 8  
Universe of Longitudinally Consistent  
Patterns Output from LSA.EXE

PERFECT LONGITUDINAL PATTERNS FOR GIVEN NUMBER OF ITEMS AND WAVES		
N PASSED	SEQUENCE	PATTERN
0	1	000 000
1	2	100 000
2	3	100 100
2	4	110 000
3	5	110 100
3	6	111 000
4	7	110 110
4	8	111 100
5	9	111 110
6	10	111 111

Table 9  
Substantive Example Illustrating an  
Absence of Longitudinal Transitions

	Time 1			Time 2		
	Low	Medium	High	Low	Medium	High
0	0	0	0	0	0	0
1	0	0	0	1	0	0
1	1	0	0	1	1	0
1	1	1	1	1	1	1
1	0	1	1	1	0	1
0	1	1	1	0	1	1

Note—Time 1 = entry into kindergarten, Time 2 = beginning of first grade. Three levels of achievement are defined: low, medium, high.

a direct inference about temporal order. Demarcating stages or levels of involvement requires an analytic model that is appropriate for dynamic processes.

The LSA analysis method described here adheres to the dynamic principles of the longitudinal Guttman simplex developed by Collins et al. (1988). Unlike cross-sectional models of behavior, which require cross-sectional item

variance, these longitudinal methods evaluate consistency across individuals in transitions along a hypothesized hierarchy of behavior. Because transitions over time are explicitly modeled, this dynamic method more accurately reflects longitudinal processes than do cross-sectional analytic methods.

In the special case where no longitudinal transitions occur (i.e., the cross-sectional hierarchy among items contains all the information), the LCS index is not simply the average of the cross-sectional scalability coefficients. In general, the LCS value will exceed the average of the CS values because longitudinal data offer greater flexibility in identifying target response patterns that minimize scalability errors. For example, LCS = 0.62 for the example data shown in Tables 5-8 while CS = 0.50 for both waves. Table 9 gives a diagram of a substantive illustration consistent with this example.

Collins and Dent (1986) note that the CL index "tends to be lower than other consistency indices. A rule of thumb is that a CL value of .30 or above represents moderate scale consistency, and anything above .5 represents very good consistency" (p. 8). Despite these helpful guidelines, it is difficult to use the CL index to make decisions about scalability. The LCS index we propose may be easier to interpret, because of the extensive information about its cross-sectional counterpart, the coefficient of scalability, CS. To evaluate their relative utilities, Monte Carlo simulations and actual data applications comparing the two longitudinal indices are needed.

#### Availability

LSA.EXE will be submitted to the National Collegiate Software Clearinghouse (Duke University Press) for distribution.

#### REFERENCES

- BAYLEY, N. (1955). On the growth of intelligence. *American Psychologist*, *10*, 805-818.
- CLIFF, N. (1979). Test theory without true scores? *Psychometrika*, *44*, 373-393.
- COLLINS, L. M., & CLIFF, N. (in press). Using the Longitudinal Guttman Simplex as a basis for measuring growth. *Psychological Bulletin*.
- COLLINS, L. M., CLIFF, N., & DENT, C. W. (1988). The Longitudinal Guttman Simplex: A new methodology for measurement of dynamic concepts in longitudinal panel studies. *Applied Psychological Measurement*, *12*, 217-230.
- COLLINS, L. M., & DENT, C. W. (1986). *LGSINDEX user's guide*. (Tech. Rep. 86-100). Los Angeles: University of Southern California, Health Behavior Research Institute.
- DOTSON, L. E., & SUMMERS, G. F. (1970). Elaboration of Guttman scaling techniques. In G. F. Summers (Ed.), *Attitude measurement* (pp. 203-213). Chicago: Rand McNally.
- GUTTMAN, L. (1944). A basis for scaling qualitative data. *American Sociological Review*, *9*, 139-150.
- KANDEL, D. (1975). Stages in adolescent involvement in drug use. *Science*, *190*, 912-914.
- KAPLAN, R. M., & SACCUZZO, D. P. (1989). *Psychological testing: Principles, applications, and issues*. Pacific Grove, CA: Brooks/Cole.
- MENZEL, H. (1953). A new coefficient for scalogram analysis. *Public Opinion Quarterly*, *17*, 268-280.



- SCHWARTZ, J. E. (1986). A general reliability model for categorical data applied to Guttman scales and current status data. In N. B. Tuma (Ed.), *Sociological methodology* (pp. 79-119). San Francisco: Jossey-Bass.
- STEWART, A. L., WARE, J. E., & BROOK, R. H. (1981). Advances in the measurement of functional status: Construction of aggregate indexes. *Medical Care*, **19**, 473-488.
- WALKER, L. J., DE VRIES, B., & BICHARD, S. L. (1984). The hierarchical nature of stages of moral development. *Developmental Psychology*, **20**, 960-966.

#### NOTE

1. As an alternative to narrowing down the potential expected patterns on the basis of the total score, LSA.EXE can compare each score with all longitudinally consistent patterns to identify the pattern that is least different. This alternative procedure is much more computationally intensive and will yield scaling coefficients (reproducibility and scalability) that are as large as or larger than those obtained from the standard method. It is also less consistent with scoring used for cross-sectional Guttman scalogram analysis and more likely than the first method to yield estimates of scalability that are too high.





