DOCUMENT RESUME

ED 216 908 SE 037 893

AUTHOR Farnsworth, Carolyn H.; Mayer, Victor J.
TITLE An Assessment of the Validity and Precision of the Intensive Time-Series Design Through Monitoring Learning Differences in Groups of Students with Formal and with Concrete Cognitive Tendencies.
PUB DATE Apr 82
NOTE 20p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; Cognitive Development; Cognitive Processes; Comprehension; Concept Formation; *Developmental Stages; Earth Science; Grade 8; Junior High School Students; *Research Methodology; Science Education; Secondary Education; *Secondary School Science; Student Characteristics; Trend Analysis
IDENTIFIERS *Science Education Research; *Time Series Analysis

ABSTRACT

This study examined whether the intensive time-series design would yield data discriminating between concrete and formal operational students and whether achievement curves exhibited in previous studies would be found with students in this study. Daily measurement of variables is characteristic of the data collection procedure in the time-series design, where individual student data are collapsed into a group mean and treated during analysis as a single subject. Data were collected for 56 school days from 95 earth science students studying a unit on plate tetonics. To determine if the design would discriminate between formal and concrete operational students (identified by Lawson's Test of Formal Reasoning), each student responded daily to a single multiple-choice item (obtained from a pool of 78 items) measuring knowledge or understanding achievement. Two parallel forms (KR20=0.80) of a multiple-choice achievement test (consisting of 45 items) were also designed from the pool of 78 items. General knowledge and understanding achievement scores for each day were obtained and analyzed, indicating the precision of the design to discriminate between students differing in cognitive ability and supporting its use in monitoring the daily acquisition of knowledge related to a concept. Results also support the validity of the design for obtaining achievement data. (JN)

# AN ASSESSMENT OF THE VALIDITY AND PRECISION OF THE INTENSIVE TIME-SERIES DESIGN THROUGH MONITORING LEARNING DIFFERENCES IN GROUPS OF STUDENTS WITH FORMAL AND WITH CONCRETE COGNITIVE TENCENCIES



By

Carolyn H. Farnsworth, Upper Arlington Public Schools

and

Victor J. Mayer, The Ohio State University

The few reported applications of the intensive time-series design in education have indicated its apparent suitability for monitoring student achievement and attitudes in the classroom (Mayer and Kozlow, 1980; Mayer and Lewis, 1979; Rojas, 1979). Daily measurement of one or more variables is characteristic of the data collection procedure of the design. Despite testing effects reported in the literature, this frequency testing does not seem to affect student achievement or produce negative attitudes toward class (Rojas, 1979). The results of the Mayer and Kozlow (1980) and the Rojas (1979) studies indicated an apparent sensitivity of the time-series design for monitoring changes in learning.

The data from the Mayer and Kozlow (1980) study showed a statistically significant trend during the intervention and a "momentum effect" in the follow-up. The intervention data of the Rojas (1979) study exhibited a statistically significant positive change in the level of the achievement data. The statistically significant trends during intervention and changes in level revealed in these two studies support the validity of an intensive time-series design.

In such designs, individual student data are collapsed into a group mean and treated during analysis as a single subject. Generalization is thus limited and replication in similar and in different situations is essential. Further evidence must also be collected to confirm the precision of the time-series design and add confidence in its validity. These are the purposes of this study.

The following questions were asked. Could the intensive time-series design yield significantly different sets of data from two groups of students who from previously reported research studies should have different learning characteristics? If it could, this would indicate the precision of the design, or its ability to discriminate between two different groups. The two groups selected were eighth grade students with formal cognitive tendencies and eighth grade students with concrete cognitive tendencies. The second question deals with the replication of results from previous studies of achievement using this design. Can the learning curves exhibited in earlier studies be found with the students and teacher in this study? If so, they will add confidence to the validity of the intensive time-series design. The dependent variable is achievement in plate tectonics and the independent variable (or intervention) is a unit on plate tectonics.

Population

Data were collected for 56 school days during December 1979 into March 1980 from 95 eighth grade earth science students. These students were in four classes meeting daily with the same teacher who was also the principal investigator of this study. The junior high school these students attended is in a middle to upper socioeconomic suburban community of Columbus, Ohio.

## Design

A multiple-group single intervention time-series design (Glass, Willson and Gottman, 1975) was adapted to the collection of daily data on achievement using as the topic of intervention, a unit on plate tectonics. Single multiple-choice items were randomly assigned to each of three groups of students, identified on the basis of their ranking on a written test of cognitive level (Lawson, 1978). The top third, those with formal cognitive tendencies, were compared on the bases of knowledge achievement and of understanding achievement with the lowest third of the students, those with concrete cognitive tendencies, to determine whether the data collected in the design would discriminate between the two groups, a test of its precision. Several studies (Goodstein and Howe, 1978; Lawson and Renner, 1975) indicated that students with formal cognitive tendencies should learn a formal concept such as plate tectonics with greater understanding than those with concrete cognitive tendencies. Data were collected daily during each of three stages: an 11-day baseline, a 30-day intervention, and a 15-day follow-up period.

## Instruments and Data Collection Procedures

Each student in the formal cognitive tendency group was matched with a student in each of the other two cognitive tendency groups, resulting in 33 sets of paired students.

Each student responded daily to a single multiple-choice achievement item and to two attitude concepts, each with different bipolar adjectives. The attitude data will be reported elsewhere. The achievement item pool consisted of 78 multiple-choice items developed to measure achievement performance at two taxonomic levels (knowledge and understanding) on the theory of plate tectonics. Each set of three students received one item each day, randomly drawn from the item pool which was stratified by taxonomic categories and instructional objectives. No set of three students received the same achievement item twice.

Two parallel forms of a multiple-item achievement instrument were designed from the pool of 78 items. Each form consisted of 45 items, twelve of which were common to both forms. Day 41, the last day of intervention, was selected for administration of the two forms. An item analysis was conducted to determine the K-R 20 reliability of each form and to determine the relative difficulty and the discrimination index of each item. The results in Table 1 indicate that the multiple-item achievement instrument had a reliability in excess of .80. Separate analysis of the 12 common items indicate no differences in the two groups taking Form A or Form B. The individual item difficulties from Forms A and B were then calculated and average item difficulty indices were used in adjusting the daily achievement data prior to analysis.

INSERT TABLE 1 HERE

Discussion of Results

The percentage of items responded to correctly by students in each of the two cognitive tendency groups was reduced to three daily achievement scores: a general achievement score, one in knowledge, and one in understanding. Product-scores were obtained for each day by multiplying the daily general achievement score, the knowledge subset, and the understanding subset by the respective mean item difficulty for that day. The product-scores for each day were plotted by cognitive tendency group. A visual inspection of the resulting profiles in Figures 1 and 2 reveals typical learning curves, an upward drift during intervention and a leveling off or drift downward during follow-up. This typical pattern of the learning curve of the product-scores provided evidence that the daily collection of data resulted in valid data and that the time-series design was sensitive enough to detect the changes in learning. Also noticeable, however, are wide daily fluctuations in the achievement data.

INSERT FIGURE 1 HERE

INSERT FIGURE 2 HERE

Two sets of analyses were performed on the achievement data. The first used linear regression techniques and t-tests to determine differences in learning between the two groups, those with formal cognitive tendencies and those with concrete cognitive tendencies. The second set of analyses was performed to examine trends within each of the two groups. The analysis used was the AUTOREG procedure of SAS, a time-series analysis program.

Tests for Differences Between Groups

Differences in the slopes of regression lines for overall achievement data were determined between the formal tendency group and the concrete tendency group for each stage of the study. The results in Table 2

INSERT TABLE 2 HERE

indicate significant differences at the .0001 level for intervention and follow-up stages and for the entire period of the study. There were no significant differences in slopes during the baseline stage. T-tests for differences between the two groups were performed on the means of the data summed across each of the three stages. Significant differences were found at the .0005 level for each of the three stages. However, the means themselves (Table 3) exhibited a small difference in

INSERT TABLE 3 HERE

baseline and much larger numerical differences in both intervention and follow-up indicated educationally significant differences in learning between the two groups.

Similar comparisons were made, based on knowledge items only and on understanding items only. Again, there were differences in slope between the two groups, significant at the .0001 level or beyond during intervention, follow-up and the entire period for both knowledge and understanding items. T-tests also were significant at the .0005 level for each stage and each of the two levels of items. Differences between means, however, were much less in the baseline than in each of the other two stages.

These results indicate that the intensive time-series design is indeed sufficiently precise in the data it yields to discriminate in learning patterns between two groups of students differing in cognitive level.

Time-Series Analysis

To examine the data for the presence of a learning curve, daily scores were subjected to a time-series analysis program (the AUTOREG program of SAS), since such a program accounts for autocorrelation within series data. Two models were specified in the analysis for each set of data, a no trend model and a model in which a positive trend was specified in the intervention stage. The trend model seemed to explain the data from the formal tendency group better than did the no trend model, accounting for 67 percent of the variance for overall achievement, 62 percent for knowledge level items, and 34 percent for understanding level items (Table 4). The trend in the overall achievement was significant at the .0001 level as was that for knowledge level

INSERT TABLE 4 HERE

items. The results of the analyses were not quite as clear cut for the concrete tendency group (Table 5), with only 28 percent of the variance

INSERT TABLE 5 HERE

accounted for in the overall achievement and less for each of the sub-groups of items. The trend model did account for slightly more variance than did the no trend model. Only the overall achievement trend, however, was significant at the .03 level.

These analyses add confidence to the validity of the intensive time-series design, since a strong indication of a learning curve is present in the data from the formal cognitive tendency group. The fact that there is no such strong indication of a trend with the data from the concrete cognitive tendency group further supports the expected results, since a number of studies (Goodstein and Howe, 1978; Lawson and Renner, 1975) have indicated that concrete level students have difficulty learning and understanding formal or abstract concepts like the concept of plate tectonics.

Conclusions

Statistically significant differences were found between the two cognitive groups both in the slopes of regression lines (.0001) and in t-tests (.0005) on the knowledge and the understanding levels of learning. These results confirm the precision of the intensive time-series design and indicate that it can be used to distinguish differences in learning between students having formal cognitive tendencies and those having concrete cognitive tendencies.

The time-series analysis using a model having a trend in the intervention was preferable to a model with no trend for both groups of students in that it accounted for a greater amount of variance in the data on knowledge and understanding levels of learning. This finding adds additional confidence to the validity of the design for obtaining achievement data. The analysis model using a trend applied to the data from the group with formal cognitive tendencies accounted for a greater degree of variance than the same model applied to the data from the group with concrete cognitive tendencies. This more conservative analysis, therefore, gave results consistent with those from the linear regressive techniques and _t_-tests, further supporting confidence in the precision of the design.

The results of this study add to the confidence that can be placed in the intensive time-series design as shown from its use in previous studies. In addition, its precision or ability to discriminate between the learning patterns of two different, though related, groups has been demonstrated. It now appears that this design can be quite useful for the investigator in concept development in that it permits the daily monitoring of the acquisition of knowledge relating to a concept.

Additional work on the design now needs to be concentrated in the suppression and/or explanation of day to day fluctuation in achievement and in adapting it to obtain types of data other than achievement data.

## Selected Bibliography

Farnsworth, C. H. "Using an intensive time-series design to develop profiles of daily achievement and attitudes of eighth-grade earth-science students at different cognitive levels during the study of the theory of plate tectonics." Unpublished doctoral dissertation. The Ohio State University, Columbus, Ohio, 1981.

Glass, G. V., Willson, V. L. and Gottman, J. M. Design and analysis of time-series experiments, Boulder, Colorado: Colorado Associated Press, 1975.

Goodstein, M., and Howe, A. C. The use of concrete methods in secondary chemistry instruction. Journal of Research in Science Education, 1978, 15 (5), 361-366.

Lawson, A. E. The development and validation of a classroom test of formal reasoning. Journal of Research in Science Teaching, 1978, 15 (1), 11-24.

Lawson, A. E., and Renner, J. W. Relationships of science subject matter and development levels of learners. Journal of Research in Science Teaching, 1975, 12 (4), 347-358.

Mayer, V. J., and Kozlow, M. J. An evaluation of a time-series single-subject design used in an intensive study of concept understanding, Journal of Research in Science Teaching, 1980, 17, 455-461.

Mayer, V. J., and Lewis, D. K. An evaluation of a time-series single-subject design. Journal of Research in Science Teaching, 1979, 16, 137-144.

Rojas, C. A. "A study of the effect of frequency of testing upon the measurement of student attitudes toward science class and measurement of achievement on a crustal evolution unit using a time-series design." Unpublished doctoral dissertation. The Ohio State University, Columbus, Ohio, 1979.

TABLE 1

SUMMARY STATISTICS OF ITEM ANALYSIS

OF MULTIPLE-ITEM ACHIEVEMENT INSTRUMENT, DAY 41

| | Form A | Form B |
|---|---|---|
| Number of Students | 51 | 44 |
| Number of Items | 45 | 45 |
| Mean | 32.39 | 32.23 |
| Standard Deviation | 7.96 | 7.32 |
| K-R 20 | 0.89 | 0.86 |
| Mean Item Difficulty | 0.28 | 0.28 |
| Mean Item Discrimination | 0.42 | 0.42 |

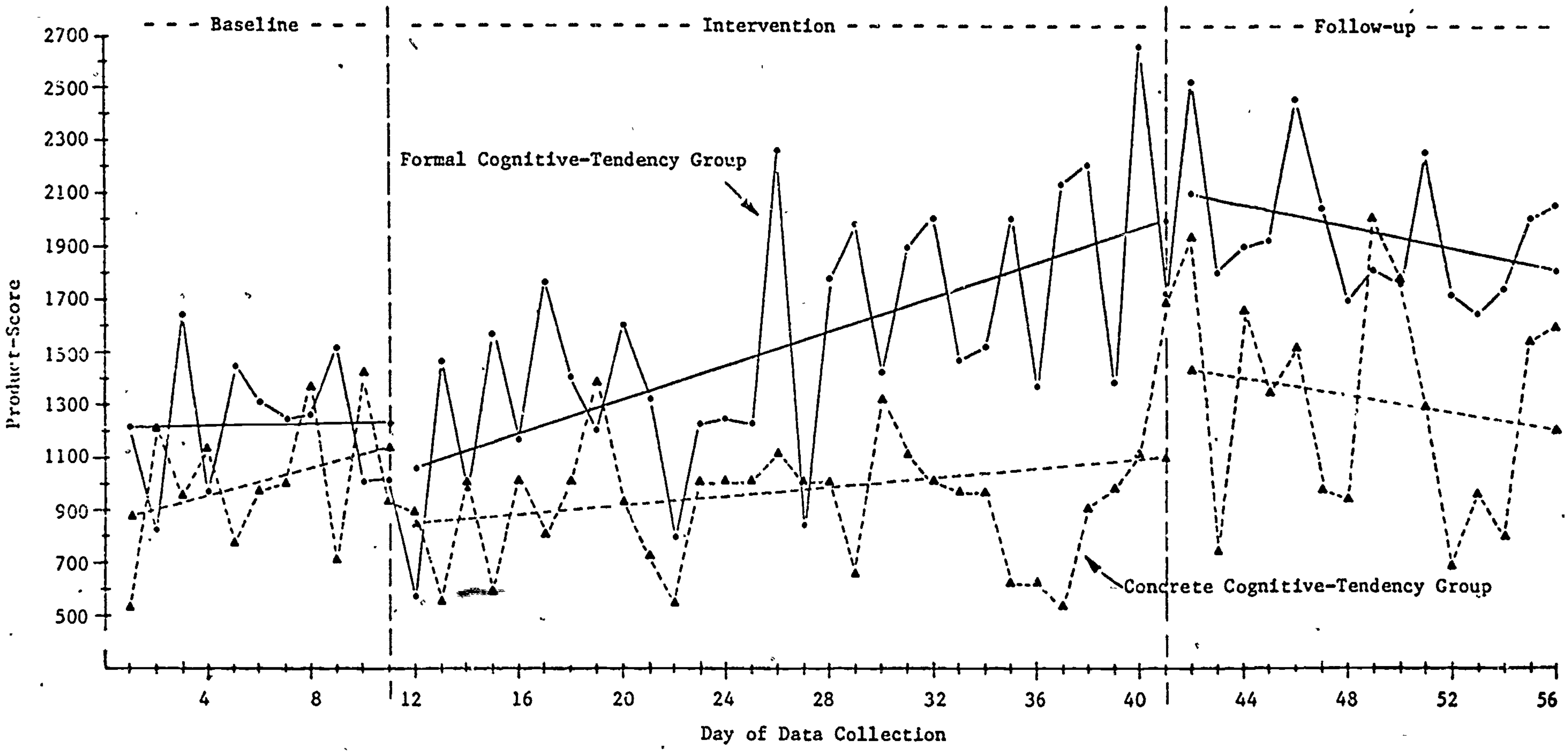


Figure 1. Profiles of Daily Product-Scores of Knowledge Items by Cognitive Tendency Group.

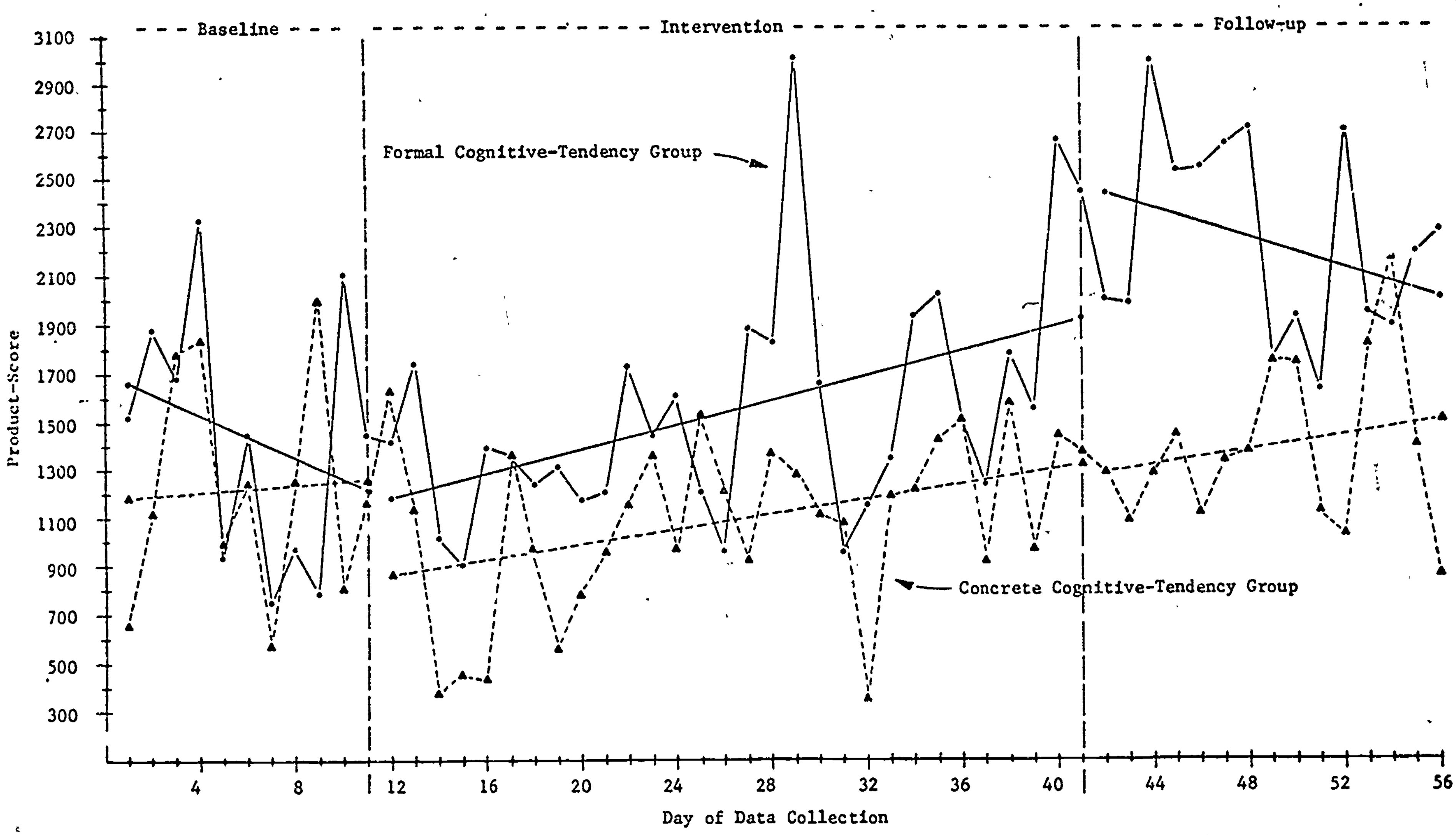


Figure 2. Profiles of Daily Product-Scores of Understanding Items by Cognitive Tendency Group.

TABLE 2

COMPARISON OF LINEAR REGRESSION LINES

ON DAILY PRODUCT-SCORES BETWEEN COGNITIVE TENDENCY GROUPS

| | Total Period $F$ (4, 104) | Baseline $F$ (2, 28) | Intervention $F$ (2, 54) | Follow-up $F$ (2, 26) |
|---|---|---|---|---|
| Achievement | 28.37* | 2.66 | 32.15* | 24.79* |
| Knowledge | 18.56* | 2.17 | 26.68* | 18.56* |
| Understanding | 12.20* | 0.70 | 9.57* | 20.04* |

* $p < .0001$

TABLE 3

DIFFERENCE IN MEANS OF

ACHIEVEMENT PRODUCT-SCORES BETWEEN COGNITIVE TENDENCY GROUPS

| | Formal Tendency | | Concrete Tendency | | |
|---|---|---|---|---|---|
| | M | SD | M | SD | t |
| Achievement | | | | | |
| Baseline[a] | 1321.40 | 204.50 | 1110.54 | 274.37 | 17.56* |
| Intervention[b] | 1559.94 | 301.85 | 1038.71 | 209.28 | 53.96* |
| Follow-up[c] | 2131.74 | 289.71 | 1367.43 | 310.49 | 61.47* |
| Knowledge Subset | | | | | |
| Baseline | 1230.68 | 256.07 | 1007.74 | 274.06 | 17.64* |
| Intervention | 1520.63 | 381.03 | 978.98 | 262.60 | 49.97* |
| Follow-up | 1956.69 | 264.21 | 1317.20 | 454.04 | 49.95* |
| Understanding Subset | | | | | |
| Baseline | 1445.56 | 547.34 | 1229.22 | 525.55 | 12.03* |
| Intervention | 1558.20 | 456.59 | 1094.88 | 355.19 | 38.07* |
| Follow-up | 2263.36 | 407.80 | 1394.31 | 358.67 | 40.27* |

[a] Degrees of freedom for baseline comparisons = 9.

[b] Degrees of freedom for intervention comparisons = 28.

[c] Degrees of freedom for follow-up comparisons = 13.

* $p < .0005$

TABLE 4

COMPARISON OF TWO ANALYSIS MODELS

OF ACHIEVEMENT DATA FOR THE FORMAL COGNITIVE-TENDENCY GROUP

| | No Trend | | | Trend | | | |
|---|---|---|---|---|---|---|---|
| | $\underline{F}$ (2, 54) | $\underline{p}$ | R-square | $\underline{F}$ (3, 54) | $\underline{p}$ | R-square | $\underline{p}_T$ [a] |
| Achievement | 15.06 | .0001 | .36 | 35.01 | .0001 | .67 | .0001 |
| Knowledge | 9.67 | .0003 | .26 | 29.24 | .0001 | .62 | .0001 |
| Understanding | 8.34 | .0007 | .24 | 8.99 | .0001 | .34 | .02 |

[a] $\underline{p}_T$ = probability that the coefficient for the trend parameter in the regression equation is statistically significant.

TABLE 5

COMPARISON OF TWO ANALYSIS MODELS

OF ACHIEVEMENT DATA FOR THE CONCRETE COGNITIVE-TENDENCY GROUP

| | No Trend | | | Trend | | | |
|---|---|---|---|---|---|---|---|
| | $F_{(2, 54)}$ | $p$ | R-square | $F_{(3, 54)}$ | $p$ | R-square | $p_T$ [a] |
| Achievement | 5.70 | .006 | .17 | 6.83 | .0006 | .28 | .03 |
| Knowledge | 5.24 | .008 | .16 | 4.07 | .01 | .19 | .2 |
| Understanding | 2.22 | .1 | .08 | 3.20 | .03 | .15 | .06 |

[a] $p_T$ = probability that the coefficient for the trend parameter in the regression equation is statistically significant.