ABSTRACT
        Developmental patterns in the solution of verbal
analogies, especially the recognition of higher-order analogical
relations, were traced. The investigation sought to: (1) provide new
developmental tests of a componential theory of analogical reasoning;
(2) identify strategy changes during the transition from midchildhood
(grade 3) to adulthood (college); (3) compare the development of
reasoning across five different verbal relations; and (4) to validate
a new method (stem-splitting) of isolating component cognitive
processes. Twenty subjects in each of grades three, six, nine, and in
college were tested on their abilities to solve verbal analogies
based on synonymy, antonymy, category membership, linear ordering,
and function in three different formats. It was found that the
componential theory was successful in accounting for response-time
data at all grade levels and error data at all but the college level;
third- and sixth-graders used a different strategy from the strategy
used by ninth-graders and college students; antonymous and functional
relations were the easiest to process at all levels; and the new
method of isolating component cognitive processes provided new
insights into age-related strategy changes in the solution of verbal
analogies. (Author/MH)

Patterns in the Solution of Verbal Analogies

Robert J. Sternberg and Georgia Nigro

Department of Psychology
Yale University
New Haven, Connecticut 06520
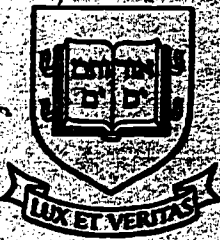
Developmental Patterns in the Solution of Verbal Analogies

Robert J. Sternberg and Georgia Nigro

Yale University

Running head: Solution of Verbal Analogies

Send proofs to Robert J. Sternberg
                Department of Psychology
                Yale University
                Box 11A Yale Station
                New Haven, Connecticut 06520

3

Abstract

This research sought to trace developmental patterns in the solution
of verbal analogies, especially with respect to the development of the abili-
ty to recognize higher-order analogical relations. Particular aims of the
investigation were (a) to provide new developmental tests of a componential
theory of analogical reasoning (Sternberg, 1977a, 1977b; Sternberg & Rifkin,
1979); (b) to identify strategy changes during the transition from mid-
childhood (grade 3) to adulthood (college); (c) to compare the development
of reasoning across five different verbal relations; and (d) to validate a
new method of isolating component cognitive processes, the "method of stem-
splitting," which can be applied to global reasoning performance.

Twenty subjects in each of grades three, six, nine, and in college were
tested in their relative abilities to solve 180 verbal analogies based on five
different verbal relations (synonymy, antonymy, category membership, linear
ordering, functional) and presented in three different formats (varying in
the number of terms in the analogy stem versus the number in the analogy options).
Both response time and error data were collected. It was found that (a) the
componential theory was successful in accounting for response-time data at
all grade levels and error data at all but the college level; (b) third-graders
and sixth-graders used a strategy that differed in key respects from the
strategy used by ninth-graders and college students; (c) antonymous and
functional relations were easier to process at all levels than were synony-
mous, category membership, and linear ordering relations; and (d) the new
method of isolating component cognitive processes provided new insights into
age-related strategy changes in the solution of verbal analogies.

Developmental Patterns in the Solution of Verbal Analogies

Reasoning by analogy is pervasive in everyday experience, and perhaps in part because of its pervasiveness, it has played a key role in psychological theories of intelligence and cognition. One of the earliest theories of intelligence, that of Spearman (1923), based its three "principles of cognition" upon three processes used in solving analogies. In an analogy of the form A is to B as C is to D, the apprehension of experience corresponds to the encoding of each analogy term, the eduction of relations to the inference of the relation between A and B, and the eduction of correlates to the application of the inferred relation from C to D.

Analogy has also played a key role in Piaget's (1949, 1950) theory of intelligence. Lunzer (1965) has argued that formal reasoning in the Piagetian sense can be identified with the recognition of second-order relations which, in reasoning by analogy, is recognition of the higher-order relation that links the relation between A and B to the relation between C and D. In essence, then, the higher-order relation constitutes the principle of analogy. Piaget's own words seem to support Lunzer's argument: Inhelder and Piaget (1958) wrote that "this notion of second-degree operations...expresses the general characteristic for formal thought" (p. 254). To the extent, then, that the stage of formal operations represents a discrete break from the stage of concrete operations, one would expect the formal-operational child (or adult), but not the concrete-operational child, to demonstrate an ability to recognize second-order relations, and to use them in solving analogies. If children who have not yet entered the formal-operational stage solve analogies, it should be by a method that is essentially non-analogical. A pronounced strategy shift should then be apparent upon the child's entrance into the stage of formal operations.

There is, in fact, an array of evidence suggesting that qualitative shifts
in strategy for solving analogies do occur as children grow older. Theorists
disagree, however, as to just what forms these shifts take, and at just what
ages they occur. Theorists can be divided roughly into two camps, those
emphasizing reasoning processes, and those emphasizing word association. This
division is only a rough one, however, since process theorists have generally
acknowledged at least some use of word association by children, and vice versa.

Consider first some theory and findings of the process theorists. Pia-
get, with Montangero and Billeter (1977), has suggested three stages in the
development of analogical reasoning. Understanding of these stages requires
some knowledge of the paradigm these investigators used to explore the de-
velopment of analogical reasoning. The investigators presented 29 children
between the ages of 5 and 13 with sets of pictures, and asked the children to
arrange the pictures in pairs. The children were then asked to put together
those pairs that went well together, placing groups of four pictures in
2 x 2 matrices that represented relations of analogy among the four pictures.
Children who had difficulty at any step of the procedure were given prompts
along the way. Children who finally succeeded were presented with a counter-
suggestion to their solution, in order to test the strength of their commit-
ment to their proposed response. At all steps along the way, children were
asked to explain the reasons for their groupings. In the first proposed
stage, characterizing the performance of children of ages 5 to 6, children
can arrange pictures into pairs, but ignore higher-order relations between
pairs. Thus, although these children can link $\underline{A}$ to $\underline{B}$ or $\underline{C}$ to $\underline{D}$, they
cannot link $\underline{A}$-$\underline{B}$ to $\underline{C}$-$\underline{D}$. In the second stage, characterizing the performance
of children from about 8 to 11 years of age, children can form analogies, but

when challenged with counter-suggestions, readily rescind their proposed

analogies. Piaget takes this as evidence of only a weak or tentative level

of higher-order reasoning ability. It is of interest, nevertheless, that

children presumably not yet at the stage of formal operations show at least

incipient analogical reasoning abilities. In the third stage, characterizing

the performance of children of ages 11 and above, children form analogies,

are able explicitly to state the conceptual bases of these analogies, and

resist counter-suggestions from the experimenter.

Lunzer (1965) presented children of ages 9 through 17+ with verbal analogies

varying in formal complexity and difficulty of content, although only the

formal-complexity variable will be considered here. The simplest analogies

had just one term missing (either $\underline{A}$, $\underline{B}$, $\underline{C}$, or $\underline{D}$), where more complex analo-

gies had two analogy terms missing (either $\underline{C}$ and $\underline{D}$, $\underline{A}$ and $\underline{B}$, $\underline{B}$ and $\underline{C}$, or

$\underline{A}$ and $\underline{D}$). Lunzer expected the analogies with single terms missing to be

easiest, those with $\underline{C}$ and $\underline{D}$ or with $\underline{A}$ and $\underline{B}$ missing to be more difficult,

and those with $\underline{B}$ and $\underline{C}$ or with $\underline{A}$ and $\underline{D}$ missing to be most difficult. The

third set should be harder than the second because there is no single rela-

tion linking the missing pair: Each term partakes of a different relation

in the analogy. This general prediction was upheld. To Lunzer's surprise,

however, children had great difficulty with even the simplest analogies un-

til about 9 years of age, and did not show highly successful performance until

the age of 11. Lunzer concluded that even the simplest analogies required

higher-order reasoning of the kind associated with formal-operational per-

formance. Lunzer's three stages of performance, of course, tie in neatly

with Piaget's.

Gallagher and Wright (Note 1, Note 2) have also investigated the pro-

cesses and strategies used in reasoning with verbal analogies, and with re-

lated results. Gallagher and Wright (Note 1) compared the relative abili-
ties of children in grades 4 and 6 to provide what they called "symmetrical"
or "asymmetrical" explanations for their analogy solutions. Symmetrical
explanations showed awareness of the higher-order relation linking A-B to
C-D. Asymmetrical explanations ignored this relation, dealing either only
with the C-D relation, or with both the A-B and C-D relations in isolation
from each other. Percentages of symmetrical responses increased with age
and were associated with level of performance on the analogies.

Levinson and Carpenter (1974) presented verbal analogies and quasi-analo-
gies to children of 9, 12, and 15 years of age. Analogies were open-ended
and of the standard form exemplified by "Bird is to air as fish is to ____."
Quasi-analogies were also open-ended, but are exemplified by the item "A bird
uses air; a fish uses ____." What is of particular interest here about the
two kinds of analogies is that the standard analogies seem to require recog-
nition of the higher-order analogical relationship, whereas the quasi-analogies
do not. One can say what a fish uses without relating it to anything used by
a bird: Solution of the quasi-analogy is facilitated by the fact that the
higher-order relation is given to the subject (in this case, use). The main
results of interest here were that 9-year olds correctly answered significantly
more quasi-analogies than standard analogies, whereas 12- and 15-year olds
showed no significant difference in performance on the two kinds of analogies.
Furthermore, performance on the analogies increased across age levels, whereas
performance on the quasi-analogies did not. These results provide further
evidence for the ability of formal-operational children, but not concrete-

operational children, to use higher-order relations in the solution of verbal analogies.

Sternberg and Rifkin (1979) investigated the development of analogical reasoning processes with two kinds of schematic-picture analogies. Of interest here are the results for analogies with perceptually integral attributes (see Garner, 1974). It was found that second-graders used a maximally self-terminating strategy, encoding and comparing the minimum numbers of attributes necessary to solve the problems; fourth-graders encoded terms exhaustively, but compared attributes in a self-terminating fashion. Sixth-graders and adults encoded terms exhaustively and performed early comparisons (inference of the relation between $\underline{A}$ and $\underline{B}$) exhaustively, but performed later comparisons with self-termination. This tendency to become more nearly exhaustive in information processing appears to be a general higher-order strategy in cognitive development (Brown & DeLoache, 1978), and appears to be associated with dramatic decreases in error rate over age (Sternberg, 1977b; Sternberg & Rifkin, 1979). Moreover, it was found that although fourth-graders, sixth-graders, and adults solved the analogies by mapping the higher-order relation between the two halves of the analogies, second-graders did not. Once again, then, we have support for a late developing ability to recognize and utilize higher-order relations.

Other investigators interested in analogical reasoning processes have studied these processes in the context of figural matrix problems (e.g., Jacobs & Vandeventer, 1971a, 1971b, 1972; Linn, 1973), but since their findings do not bear directly upon the present research, these findings will not be reviewed here.

Consider next some theory and findings of theorists emphasizing the use of association in the solution of verbal analogies. The pioneer in the study

of associative responding by children during analogy solution is Achenbach.

Achenbach (1970a, 1970b, 1971) found that children in the intermediate and

early secondary school grades differ widely in the extent to which they use

word association as a means of choosing one from among several response

options. Moreover, the extent to which children use word association

serves as a moderator variable in predicting classroom performance: Cor-

relations between performance on IQ tests and school achievement were

substantially lower for children who relied heavily on free association

to solve analogies than for children who apparently relied primarily on

reasoning processes. Gentile, Tedesco-Stratton, Davis, Lund, and Agunanne

(1977) further investigated children's associative responding, using

Achenbach's test. They found that associative priming can have a marked

effect on test scores, leading children either toward or away from correct solutions.

The present research uses a form of componential analysis (Sternberg, 1977b,

1978, 1979) to provide converging operations to test the stages of analogy solu-

tion proposed by Piaget, and to integrate into a single theoretical account reason-

ing and associative theories of analogy solution. More specifically, the research

is aimed at (a) providing new developmental tests of a componential theory of ana-

logical reasoning (Sternberg, 1977a, 1977b; Sternberg & Rifkin, 1979), and at

exploring the relations of the theory to Piaget's notions regarding developmental

stages on the one hand, and associative notions of analogy solution on the other;

(b) identifying strategy changes during the transition from mid-childhood (grade

3) to adulthood (college); (c) comparing the development of reasoning across

five different verbal relations; and (d) validating a new method of isolating

component cognitive processes from global reasoning performance. In order to

understand the research, the reader needs first to know something about the com-

ponential theory of analogical reasoning underlying it.

Componential Theory of Analogical Reasoning

Component Processes

Reasoning components. According to the componential theory of analogi-
cal reasoning (Sternberg, 1977a, 1977b), as many as six component processes
are used in reasoning by analogy. Consider, for example, the analogy,
HAPPY : SAD :: COLD : (SNOW, COOL, HOT, JOYFUL). According to the theory,
the reasoner (a) encodes each term of the analogy, retrieving from semantic
memory a list of attributes for each term that might be relevant to analogy
solution; (b) infers the relation between HAPPY and SAD, recognizing that
the two words are antonyms; (c) maps the higher-order relation linking the
first half of the analogy (starting with HAPPY) to the second half of the
analogy (starting with COLD), realizing that the relation of antonymy will
have to be carried over from the first half to the second; (d) applies the
previously inferred relation from COLD to at least some of the answer op-
tions, seeking an option that is antonymous to COLD; (e) optionally, if none
of the answer options seems antonymous, justifies one option as preferred to
the others, although nonideal; and (f) responds by communicating the chosen
answer, which, in this analogy, should be HOT.

Association component. The reasoning components described above have
provided a sufficient account of analogical reasoning processes used by
adults solving a variety of different types of analogies, including ones
based upon schematic pictures, words, and geometric forms. The research of
Achenbach cited above, however, makes it clear that many children use word
association in addition to or instead of certain reasoning processes. The
list of reasoning components posited by the componential theory, therefore,
must be augmented by the addition of an association component that takes
into account this associative process. The association component is assumed

to be used to link the last term of the analogy stem (COLD in the example)
to each of the various answer options. In the present view, high associa-
tion can either facilitate performance of the reasoning component linking
the analogy stem to the answer options (in the example, application), or else
replace that component altogether, resulting in the selection of the option
of highest associative value with respect to the last term in the item stem.

Alternative Strategies for Combining Component Processes

The component processes described above can be combined in several dif-
ferent ways to form an overall strategy or model for analogy solution. Dis-
cussions of strategies in previous reports on the componential theory of ana-
logical reasoning (e.g., Sternberg, 1977a, 1977b; Sternberg & Rifkin, 1979)
have dealt with alternative strategies for encoding single analogy terms and
for comparing attributes of pairs of analogy terms. These strategies or
models have been identified by means of roman numerals (I, II, III, IV).
In the present discussion, the unit of analysis will be the whole analogy term
rather than an attribute of an analogy term. In order to distinguish these
newly proposed strategy models from the ones proposed earlier, the new ones
will be identified by means of letters (A, B, C, and later, D).

Model A. In Model A, subjects scan all answer options exhaustively. In
other words, they always test all of the answer options. In the example, they
encode the first two analogy terms (HAPPY, SAD), infer the relation between
them, encode the third analogy term (COLD), map the higher-order relation
from the first half of the analogy to the second, encode the answer options,
and apply the inferred relation from the third analogy term (COLD) to each
of the answer options (SNOW, COOL, HOT, JOYFUL). Finally, they respond.
Justification would be used if none of the options were deemed ideal. To the

extent that association plays any role at all, it facilitates the linking

of the analogy stem to the options that are highly associated to the last term

of the stem (COLD).

Model B. In Model B, subjects scan answer options in an ordered,

self-terminating fashion. In other words, they test answer options top-

down until they arrive at an answer option that satisfies their criterion

for correctness. They then respond without scanning the remaining options.

In the example, their solution of the analogy would be identical to that

suggested by Model A until application, at which time they would apply the

relation of antonymy from COLD to each of SNOW, COOL, and HOT, and finding

HOT to be antonymous to COLD, would respond without checking the last option,

JOYFUL. Justification can be used in this model to decide whether a given

nonideal option satisfies the criterion for correctness, that is, is good enough.

Association facilitates linking of the analogy stem to those options that are scanned.

Model C. In Model C, as in Model B, subjects scan answer options in

self-terminating fashion. In this model, however, the order in which the op-

tions are scanned is guided by the level of association between the last term

of the stem and each of the answer options. Specifically, options are scanned

in decreasing order of associative relatedness to the last stem term, which is

assumed to be computed during the prior encoding operations. Scanning ceases

when an answer option is found that satisfies a subject's criterion for

correctness. Otherwise, processing is identical to that in Model B. Sup-

pose, in the example analogy, the highest associate of COLD is COOL, followed

by HOT, SNOW, and JOYFUL. Then a subject would scan two answer options

(as opposed to three in Model B, and four in Model A) to reach a solution.

Note that in this model the subject does not need to encode each answer option

fully. He or she need only encode the associative relatedness of each option

13

to the last term in the analogy stem.

To summarize, a theory of analogical reasoning has been proposed that

has six reasoning components and one associative component, each of which

is assumed to act upon semantic attributes of verbal analogy terms. Three

alternative strategy models have been proposed that can be used to combine

the components. A major goal of the experiment to be described is to test the

applicability of the proposed theory and alternative models to analogy so-

lution as performed by children of differing ages.

Method

Subjects

Subjects were 20 students in each of grade 3, grade 6, grade 9, and

college. Mean ages at each grade level were 8½, 11½, 14½, and 19 years.

Elementary and secondary students were from a middle-class suburb of New

Haven; college students were Yale undergraduates.[1] Subjects at each grade

level were approximately equally divided between sexes.

Materials

All subjects received the same 180 verbal analogy items. Vocabulary

level was restricted to grade three or below according to the Thorndike-

Lorge norms, so that reasoning rather than vocabulary would be the primary

key to performance. The 180 items were cross-classified in two different ways.

Of the 180 items, 36 were classified into each of the following semantic

relations: synonym (e.g., UNDER ::BENEATH :: PAIN : (PLEASURE, DOCTOR, FEELING,

HURT));antonym (e.g., START : FINISH :: FAR : (NEAR, AWAY, TRAVEL, ᵀARTHER));

functional (e.g., SHOES : FEET :: HAT : (HEAD, BUCKET, CLOTHES, CAP)); linear

ordering (e.g., YESTERDAY : TODAY :: BEFORE : (NOW, WHEN, AFTER, TIME)); and

category membership (e.g., NOON : TIME :: WEST : (DIRECTION, SUNSET, EAST,

NORTHWEST)).

Crossed with this classification were 60 items presented in each of three formats. The formats differed in the relative numbers of terms in the analogy stem versus in the analogy options. Specifically, the number of terms in the analogy stem could be either three, two, or one. The remaining terms were in the options. Consider an example of each format:

1. NARROW : WIDE :: QUESTION : (TRIAL, STATEMENT, ANSWER, ASK)

2. WIN : LOSE :: (DISLIKE : HATE)
                 (EAR : HEAR)
                 (ENJOY : LIKE)
                 (ABOVE : BELOW)

3. WEAK : (SICK :: CIRCLE : SHAPE)
          (STRONG :: POOR : RICH)
          (SMALL :: GARDEN : GROW)
          (HEALTH :: SOLID : FIRM)

Numbers of answer options varied from two to four, and were equally represented across semantic relations and item formats. Furthermore, the answer options were balanced over the five verbal relations, so that a distractor based upon any of the four incorrect semantic relations for a given problem was equally likely to appear in any item.

Procedure

Experimental procedure was the same for all subjects, except that adults completed the procedure in one long session, whereas third, sixth, and ninth graders completed the procedure in two short sessions. These two sessions occurred no more than one week apart. Testing was individual, taking place in the schools for grade-school children, and in a laboratory at Yale for adult college students.

All subjects began the first session with an introduction to the concept of a verbal analogy. Subjects were then introduced to the three different formats for analogies. After subjects received three practice items, the experi-

mental apparatus, a portable tachistoscope with attached centisecond clock,

was explained to them.   The introduction to the experiment ended

with subjects receiving 15 practice items—one practice item with each of

the five semantic relations crossed with each of the three formats.

Test items were typed in large capital letters (IBM ORATOR typeface)

on white cards and presented to subjects individually via the portable ta-

chistoscope. Response times were recorded to the nearest centisecond, and

errors in responses were noted.

In addition to the subjects in the main part of the experiment, two

other groups of subjects supplied ratings of particular aspects of the analo-

gies. These ratings were used to estimate the justification and association

parameters, as will be explained subsequently.

Twenty-four subjects compared the $\underline{A}$ to $\underline{B}$ relationship to the $\underline{C}$ to $\underline{D}$-Keyed

relationship for each analogy, rating the closeness (or degree of higher-

order relationship) between the two (lower-order) relationships.[2] Consider,

for example, the analogy NARROW : WIDE :: QUESTION : (TRIAL, STATEMENT, ANSWER,

ASK). Subjects in this group would rate the closeness of the higher-order

relationship between NARROW and WIDE on the one hand, and QUESTION and ANSWER

on the other. Ratings were made on a scale from 0 (indicating identity be-

tween the two lower-order relations) to 5 (indicating extreme disparity be-

tween the two lower-order relations). Items to be rated were presented in

booklets to small groups of adult subjects, and subjects had unlimited time

to supply the ratings.

Twenty-four other subjects rated the associative relatedness between

the last term of the analogy stem and the first term of the keyed answer op-

tion.[3] The exact pair of terms to be rated depended upon item format. In

the first format, ratings were between the third term of the stem and the keyed option (in the example above, QUESTION to ANSWER). In the second format, ratings were between the second term of the stem and the first term of the keyed option (in the example above, LOSE to ABOVE). In the third format, ratings were between the first term of the stem and the first term of the keyed option (in the example above, WEAK to STRONG). Ratings were made on a scale from 1 (indicating extremely low association) to 5 (indicating extremely high association). Items to be rated were again presented in booklets to small groups of adult subjects, and subjects had unlimited time to supply the ratings.

Design

The basic design of the experiment consisted of the crossing of 20 subjects at each of four different age levels with verbal relations and item formats. These latter two variables were both within-subjects. There were two dependent variables--response time and error rate. Independent variables used in mathematical modeling of the dependent variables are described below.

Model Testing and Parameter Estimation

The paradigm used in the present experiment is "componential" in the sense that it uses a method of task decomposition to isolate individual components of information processing. The method of stem-splitting used here is one of several methods of task decomposition that can be used for this purpose (see Sternberg, 1978). Although each analogy in the experiment requires use of the same components, the numbers of executions of various components differs across items, depending upon the format of the item and the number of answer options in the item. Different models make different predictions regarding the exact number of executions of each component needed to solve a given item. Consider, for example, an item presented in the first for-

mat, such as example 1 presented earlier. According to the fully exhaustive
Model A, subjects must encode seven analogy terms (three in the stem and
four answer options), infer one relation (between NARROW and WIDE), map
one relation (between the half of the analogy headed by NARROW and the half
headed by QUESTION), apply four relations (between QUESTION and each of the
four answer options), and respond once. The amount of justification required
is estimated on the basis of ratings between A-B and C-D. The less the
two halves of the analogy correspond, the more the justification that is required.
Were every keyed answer "perfect," no justification would be needed at all.
The amount of facilitation resulting from the association variable is esti-
mated on the basis of ratings between the last term of the stem and the
first term of the correct answer option. The higher the level of association,
the more the facilitation that occurs. The predictions of the ordered self-
terminating Model B differ from those of Model A in two respects: Only six
terms need to be encoded, and only three relations need to be applied. Because
the correct option is the third one, the fourth option need not be encoded,
nor need any relation be applied to it. Consider, finally, the predictions
of the associatively self-terminating Model C. If ASK has the highest associ-
ation to QUESTION, and ANSWER the second highest, then the predictions of Model
C differ from those of Model A in two respects: Only these two answer options
need to be encoded (in addition to the three analogy terms in the stem, for
a total of five terms to be encoded), and only two applications are needed,
the first from QUESTION to ASK, the second and final one from QUESTION to ANSWER.

A similar kind of analysis can be applied to items in the second and third
formats. Consider, for example, the predictions of Model A for the second
numbered example above: The subject needs to encode ten analogy terms, infer
one relation (between WIN and LOSE), map four relations (from the analogy half

headed by WIN to the alternative halves headed by DISLIKE, EAR, ENJOY, and

ABOVE), apply four relations (from DISLIKE to HATE, EAR to HEAR, ENJOY to

LIKE, and ABOVE to BELOW), and respond once. Justification and association

are estimated from ratings. Consider as a final example the predictions of

Model B for the third numbered example above, where the correct answer is

the second option. The subject needs to encode seven analogy terms (one

in the analogy stem and six in the first two analogy options), infer two

relations (from WEAK to SICK and from WEAK to STRONG), map two relations

(from the analogy half headed by WEAK to the alternative halves headed by

CIRCLE and POOR), apply two relations (from CIRCLE to SHAPE and from POOR

to RICH), and respond once. Again, justification and association are es-

timated from ratings.

This method of stem-splitting enables one to isolate each component

in the componential theory of analogical reasoning, and moreover, to dis-

tinguish between the predictions of the three alternative models (plus,

possibly, others not proposed here). Model testing and parameter estimation

were done by multiple regression. Both response time and error rate were

used as dependent variables. Independent variables were required numbers of

encodings, inferences, mappings, applications, and responses needed under

each model, plus justification and association ratings. Because these latter

two variables were not measured on an absolute scale, raw regression weights

were not interpretable. Standardized regression coefficients (beta weights)

were used to indicate the contribution of each variable to the prediction of

response time or error rate.

In the model-testing procedure, response time was hypothesized to equal

the sum of the amounts of time spent on each of the seven components. A

simple linear model predicts response time to be the sum across the different

19

components of the number of times each component is executed (as an independent variable) multiplied by the duration of its execution (as an estimated parameter). Proportion of errors is hypothesized to equal the sum of the difficulties encountered in executing each component. A simple linear model predicts proportion of errors to be the sum across the different components of the number of times each component is executed (as an independent variable) times the difficulty of its execution (as an estimated parameter). The assumptions underlying model testing are explained more fully in Sternberg (1977b).

Results

Qualitative Data Analyses

Figure 1 shows developmental patterns in response time and error rate for analogies based upon each of the five semantic relations; Figure 2 shows developmental patterns for these dependent variables for analogies presented in each of the three item formats. Because response time and error rate were correlated across item types (r = .72, .76, .67, and .18 for grade 3, grade 6, grade 9, and college levels respectively), a multivariate analysis of variance was conducted upon the response times and error rates considered jointly as dependent variables. Independent variables were grade level, semantic relation, and item format. In all cases, multivariate F was estimated using Wilks's lambda criterion (λ). Several striking patterns appeared in the data.

Insert Figures 1 and 2 about here

We will consider initially the three main effects. First, it can te seen in Figure 1 that except for two perturbations in the error data, response times and error rates decreased monotonically across grade levels for each

of the five semantic relations. The difference across grade levels was highly significant, $F(6,2278) = 133.62$, $p < .001$. Second, the plots in Figure 1 suggest that the relations were differentially difficult, a suggestion confirmed by the analysis of variance, $F(8,2278) = 21.54$, $p < .001$. In particular, functional and antonymous relations showed shorter response times and lower error rates across grade levels than did the other semantic relations, which were not well distinguished among themselves. Third, Figure 2 shows a substantial effect of item format upon both response times and error rates, an effect that is highly significant, $F(4,2278) = 27.94$, $p < .001$. Thus, each of the independent variables in the experimental design resulted in a significant main effect.

Next, consider interactions between these variables. First, the interaction between grade and semantic relation was not significant, $F(24,2278) = 1.43$, $p > .05$. This additivity of grade and semantic-relation effects is apparent in Figure 1, where response times and error rates appear to decline at roughly the same rates across grade levels for the various semantic relations. Second, the interaction between semantic relation and item format (not shown in the figures) was statistically significant, $F(16,2278) = 4.94$, $p < .001$. Third, the interaction between grade and item format was significant, $F(12,2278) = 5.57$, $p < .001$. This interaction was particularly pronounced for response times, as can be seen in Figure 2. Whereas response times increased monotonically across item formats for grade 9 and adults, response times showed a curvilinear pattern across item formats in grades 3 and 6. This difference in response-time patterns is sufficient to suggest that the strategies of third and sixth graders differ from those of ninth graders and college students, although there is no evidence of differences within each of these two groupings. Finally, the triple interaction

of grade, relation, and item format was nonsignificant, $\underline{F}(48,2278) = .93$, $\underline{p} > .05$.

Univariate analyses of variance were also conducted on response times and error rates in order to follow up on the multivariate analysis. Since the results of these univariate analyses were comparable to those of the multivariate analysis, they will not be considered separately here.

Quantitative Data Analyses

Simple correlations of model predictors with response time and error rate. Simple correlations were computed between each of the predictors of the componential theory of analogical reasoning and both response time and error rate.[4] Separate correlations were computed for each grade level for each of the predictors as conceptualized under each of the three models described earlier. Response times for erroneous responses were removed from these and subsequent data analyses.

A fourth model, Model D, was entered into consideration after inspection of the data. This model is actually a mixture of Models A and B, in that whether exhaustive or ordered self-terminating scanning of answer options is used depends upon item format. In particular, this model assumes that exhaustive scanning of options is used if items are presented in either the first or second format (three or two terms in the analogy stem), and that ordered self-terminating scanning of options is used if items are presented in the third format (one term in the analogy stem). Obviously, other mixture models are also possible. In retrospect, however, this model seems particularly plausible, because the third item format, requiring by far the largest number of encodings of analogy terms, seems to place far greater demands upon working memory than do either of the other two models. This model, therefore, might be used by subjects whose working memories were not adequate to the

demands of exhaustive processing in the third format.

Insert Table 1 about here

Consider first the usefulness of the predictors in the componential theory in accounting for response times and error rates, without regard to the particular model under this theory that subjects at a given grade level happened to use. The success of the various predictors in accounting for response times and error rates can be assessed by looking at the columns of the table. At least some successful prediction was obtained for response times generated at all grade levels, and for error rates generated at all but the college level. As it happens, Sternberg (1977a, 1977b) was also unsuccessful in predicting error rates of college students for similar analogies. It was suggested in this previous work that the few errors adults make on analogies with relatively low-level analogy terms are due to idiosyncratic knowledge gaps. Hence, a general model of errors in simple verbal analogies may not be possible at the college level. The present results are consistent with this view.

Consider next the usefulness of the individual predictors, again without regard to the particular model used. Here, we look at groups of rows in the table. It can be seen that encoding, mapping, and application were predictive of performance at all grade levels. Justification was the only variable that was not predictive at any grade level, although this failure in prediction will be shown later to hold only at the level of simple (zero-order) correlations. The most interesting results are probably those for inference and association, since the predictive power of each of these variables interacts with grade level, and in opposite ways. Inference is correlated with response time at the grade 9 and college levels, but not at the grade 3 and

and grade 6 levels. Association, on the other hand, is correlated with response time at the grade 3 and grade 6 levels, but not at the grade 9 and college levels.[5] These two results, taken together, imply a difference in strategy between third graders and sixth graders on the one hand, and ninth graders and college students on the other. This implication is consistent with the interaction between item format and grade level noted earlier, which suggested that third and sixth graders solved the verbal analogies using a strategy different from that of ninth graders and college students. It appears that children in the two earlier grades rely heavily on associative processes in analogy solution: High association between the last term of the item stem and the correct response option facilitates the speed and accuracy of the younger children. The fact that these children do not increase the amount of time spent on inference as a function of the number of inferences to be made suggests that association is used, at least to some extent, as a substitute for full reasoning by analogy. Older children appear to rely almost exclusively on reasoning processes in analogy solution: High association does not facilitate solution. Moreover, increases in the number of inferences to be made results in an increase in response latency, as would be required in any model based exclusively upon reasoning processes.

Consider finally the validity of the four componential models in predicting response times. The most useful variable in distinguishing among models is encoding, since this variable was the one most highly associated with response times. Inspection of the correlations between encoding and response time reveals that correlations were highest for Model D at the third and sixth grade levels, but for Model A at the ninth grade and college levels. Most clearly, Model D is best at the earlier two grades, but certainly not at the latter two grades. The pattern for inference is less

clear, and inference is not predictive under any model at the third and

sixth grade levels. The pattern for mapping is the same as that for encoding.

The pattern for application again, is less clear, although Model D certainly

fails at the latter two grade levels.

Similar correlations were computed separately for each semantic rela-

tion at each grade level. The results for each relation reflected those for

the combined relations, which is not surprising in view of the nonsignificance

of the interaction between grade level and semantic relation. Since these

separate correlations were not interesting, they are not presented here.

To summarize, there is at least tentative evidence of a strategy

shift between the sixth and ninth grade levels (ages 11½ and 14½). Older

children and adults appear to rely primarily upon reasoning processes in the

solution of analogies, whereas younger children appear to rely primarily upon

associative processes. Moreover, older children and adults appear to use

fully exhaustive scanning of answer options, whereas younger children appear

to use fully exhaustive scanning of answer options only in the first two

item formats. In the third format, where the demands upon working memory

are the greatest, these children appear to rely upon self-terminating scanning

of options. Self-terminating scanning reduces memory load, in that it requires

less storage of results from earlier option scans. The use of self-terminating

scanning of options in the third format results in these items being solved

more quickly than those in the second format, where full exhaustive scanning

is used.

Multiple correlations of two best model predictors with response time

and error rate. Multiple correlations were computed between the predictors

of the various models and both response time and error rate.[6] The quality

of the data and the number of degrees of freedom for residual proved suffi-

cient to allow only two predictors to be entered into each regression with

44444444444444444444444444444444444444444444444444444444444444444444444444444444444444444444444444444

confidence. Additional predictors were reliable in some cases but not in others, and seemed to follow no consistent pattern. As will be seen, the use of two predictors permitted significant increments to prediction over that obtained with just one variable in all but one case (Model B at grade 9).

Insert Table 2 about here

Table 2 presents multiple correlations for the two best predictors of response time under each model at each grade level, as well as standardized regression coefficients (beta weights) for each predictor and reliabilities. Because the various independent variables were on different scales—some variables were counts (encoding, inference, mapping, application) and others were ratings (justification and association)—raw regression coefficients were not meaningful.

Consider first the column at the far right of the table. This column shows the square root of the reliability coefficient (called the reliability index) for the response-time data at each grade level. The reliability is of the internal-consistency type (coefficient alpha). The reliability indices are of interest because they show that the data were quite reliable at each grade level, and they also show the maximum possible multiple correlation under the assumptions of classical test theory. Note that the reliabilities show an increasing trend over grade levels, in particular, between grade 6 and grade 9. This transition between grades 6 and 9 is the interval in which the previous data suggested the occurrence of a strategy change. Moreover, the increase in reliabilities over age is consistent with previous findings based upon schematic-picture analogies (Sternberg & Rifkin, 1979). The increase in reliability suggests that subjects tend to become more consistent in their use of strategy with increasing age.

Second, consider the identity of the model best accounting for the re-
sponse-time data at each grade level. This model is D at grades 3 and 6,
and A at grade 9 and college. This pattern of model support is consistent
with that obtained for the simple correlations. The value of the multiple
correlation differs from that of the reliability index by .06, .03, .07,
and .05 at each of the respective grade levels. Thus, the best model does
almost as well as any model could do under the assumptions of classical
test theory. The proximity of the multiple correlation to the reliability
index shows an additional reason why adding additional variables to the
prediction equation would not have been justified. Two variables did about
as well as any number of variables could be expected to do. Additional
variables would be more likely to capitalize upon chance fluctuations in
the data than to contribute to reliable prediction of response times.

Third, consider the identities of the variables contributing to the
best model at each grade level. As it turns out, the identical variables
enter into the multiple regression at each level--encoding and justification--
thereby making fits of the models comparable across grades as well as within
grades. Each of these variables contributes significantly and positively
at each level. The entrance of encoding into the multiple regression equa-
tion comes as no surprise, since it was the single most powerful predictor
of response times in the simple correlations. The entrance of justification
seems more surprising, since it was the one variable that showed no signifi-
cant relation   response time in the simple correlations. Entrance of the
second variable into the regression equation, however, is determined by the
identity of the variable having the highest partial correlation with the cri-
terion after variance shared with the first variable is removed from each
of the potential second variables and the criterion.[7] Here, justification

27

was always the variable with the highest partial correlation (in the best mode
at each grade level). Conceptually, this result indicates that the effect of
justification upon response time was statistically independent of, and was
masked by, the effect of varied numbers of encodings. Justification, it will
be recalled, is a function of the degree of fit, or analogy, between the
first two items of the analogy and the last two items. It is psychologically
(as well as statistically) independent of the nature and number of incorrect
answer options. When the effect of these incorrect options upon response time
was removed, the effect of justification could show itself. The strength of
justification as a predictor of response time is consistent with previous
findings. In the one previous study in which justification was estimated
(geometric analogies in Sternberg, 1977b), justification was the single most
powerful predictor of response time. Encoding was the second most powerful
predictor.

The effect of justification upon response time at the third and sixth
grade levels suggests that younger children are not wholly insensitive to the
conceptual analogy between the first two terms and the last two terms of the
analogy. An alternative explanation of this effect, however, might be that the entrance
of justification into the regression equation for the younger children was due
not to their recognition of a conceptual analogy, but to their recognition of
an associative analogy. On this account, younger children view the two halves
of an item as analogous to the extent that the level of association between the
first two terms matches that between the last two terms. Since level of associa-
tive analogy would be likely to correlate with level of conceptual analogy, justi-
fication would enter the regression equation, but spuriously. In order to test
this hypothesis, ratings were collected of the level of associative relation
between the first two terms of the analogy and between the last two terms of the

analogy. The absolute value of the difference between these two levels of association was then computed as an index of the degree of associative mismatch between the two halves of the analogy. If associative rather than conceptual mismatch was responsible for the entrance of justification into the regression equation, then the new associative justification parameter should replace the conceptual one when both are given the opportunity to enter the equation. In fact, however, the conceptual justification parameter entered the equation first, and once this parameter entered, the remaining contribution of the associative justification parameter was trivial. Thus, any contribution of the associative justification parameter was due to its overlap with the conceptual justification parameter, rather than the other way around. The conclusion to be drawn from this analysis seems to be that even third graders have some sensitivity to conceptual analogies.

Multiple correlations were also computed for the two best predictors of error rates. The error data were less useful in distinguishing among models than were the response-time data, probably in part due to their lower reliability. The multiple correlations for the models identified as best above were .67 for Model D at grade 3 (based upon encoding and association), .72 for Model D at grade 6 (based upon encoding and justification), .64 for Model A at grade 9 (based upon application and justification), and (a nonsignificant) .37 for Model A at the college level. Reliability indices of the error data at the respective grade levels were .93, .93, .92, and .79. Thus, the error models provided good, but imperfect fits to the true variation in the error data.

Discussion

The present research sought to provide a converging test of the stages of analogy solution proposed by Piaget. The outcomes of the research were mixed in their support of Piaget's model of cognitive development, and also were

mixed in their support of the three stages in Piaget's developmental model of analogical reasoning. Two stages seemed to be discernible, although it was not clear that they corresponded to any two of Piaget's stages, since there was some evidence of true reasoning even in the earlier stage. In a first stage, whose occurrence coincides roughly with the ages associated with concrete operations, solution is primarily but not exclusively associative: Association affects but does not wholly control analogy solution. Reasoning in this stage is incomplete rather than absent: Children do not infer all possible relations between possible initial pairs of analogy terms, and they terminate information processing prematurely if information ad goes beyond their working-memory capacity. In a second stage, whose occurrence coincides roughly with the ages associated with formal operations, children and adults fully relate the first half of the analogy to the second half. Solution is accomplished by verbal reasoning rather than by verbal association.

According to the proposed account of strategy development, children become more exhaustive in their information processing with increasing age. This tendency toward greater use of exhaustive information processing is consistent with previous developmental findings with People Piece analogies (Sternberg & Rifkin, 1979), and is also consistent with what appears to be a general metacognitive tendency toward the use of more nearly exhaustive strategies with increasing age (Brown & DeLoache, 1978).

In general, the results of the research provide further developmental support for the proposed componential theory of analogical reasoning (Sternberg, 1977a, 1977b; Sternberg & Rifkin, 1979). Five of the six reasoning processes postulated by the theory showed statistically significant zero-order correlations with response time, and the other process, justification, showed a statistically significant first-order (partial) correlation. It was necessary to augment the previously proposed theory by an association component, which

appears to be used by the younger children but not the older ones. The

same components and strategy appear to have been applied at each grade

level to analogies based upon each of the five semantic relations. Thus,

both the proposed theory and strategy models appear to be generalizable,

at least within the verbal domain.

Reference Notes

1. Gallagher, J. M., & Wright, R. J. Children's solution of verbal analogies: Extension of Piaget's concept of reflexive abstraction. Paper presented at the annual meeting of the Society for Research in Child Development, New Orleans, 1977.

2. Gallagher, J. M., & Wright, R. J. Piaget and the study of analogy: Structural analysis of items. Paper presented at the Eighth Annual International Interdisciplineary UAP Conference on Piagetian Theory and its Implications for the Helping Professions, University of Southern California, 1978.

References

Achenbach, T. M. Standardization of a research instrument for identifying
associative responding in children. Developmental Psychology, 1970, 2,
283-291. (a)

Achenbach, T. M. The children's associative responding test: A possible
alternative to group IQ tests. Journal of Educational Psychology,
1970, 61, 340-348. (b)

Achenbach, T. M. The children's associative responding test: A two-year
followup. Developmental Psychology, 1971, 5, 477-483.

Brown, A. L., & DeLoache, J. S. Skills, plans, and self-regulation.
In R. Siegler (Ed.), Children's thinking: What develops?
Hillsdale, N.J.: Erlbaum, 1978.

Garner, W. R. The processing of information and structure. Hillsdale, N.J.:
Erlbaum, 1974.

Gentile, J. R., Tedesco-Stratton, L, Davis, E., Land, N. J., & Agunanne, B. A.
Associative responding versus analogical reasoning by children. Intelli-
gence, 1977, 1, 369-380.

Inhelder, B., & Piaget, J. The growth of logical thinking from childhood to
adolescence. New York: Basic Books, 1958.

Jacobs, P. I., & Vandeventer, M. The learning and transfer of double-
classification skills by first graders. Child Development, 1971, 42,
149-159. (a)

Jacobs, P. I, & Vandeventer, M. The learning and transfer of double-
classification skills: A replication and extension. Journal of
Experimental Child Psychology, 1971, 12, 140-157. (b)

Jacobs, P. I., & Vandeventer, M. Evaluating the teaching of intelligence.

Psychology in the Schools, 1972, 10, 67-75.

Levinson, P. J., & Carpenter, R. L. An analysis of analogical reasoning in children. Child Development, 1974, 45, 857-861.

Linn, M. C. The role of intelligence in children's responses to instruction. Psychology in the Schools, 1973, 10, 67-75.

Lunzer, E. A. Problems of formal reasoning in test situations. In P. H. Mussen (Ed.), European research in cognitive development. Monographs of the Society for Research in Child Development, 1965, 30, 19-46.

Piaget, J. Traité de logique. Paris: Colin, 1949.

Piaget, J. The psychology of intelligence. New York: Harcourt, Brace, 1950.

Piaget, J. (with Montangero, J., & Billeter, J.). Les correlats. L'abstraction réfléchissante. Paris: Presses Universitaires de France, 1977.

Spearman, C. The nature of 'intelligence' and the principles of cognition. London: Macmillan, 1923.

Sternberg, R. J. Component processes in analogical reasoning. Psychological Review, 1977, 84, 353-378. (a)

Sternberg, R. J. Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Hillsdale, N.J.: Erlbaum, 1977. (b)

Sternberg, R. J. Isolating the components of intelligence. Intelligence, 1978, 2, 117-128.

Sternberg, R. J. The nature of mental abilities. American Psychologist, 1979, 34, 214-230.

Sternberg, R. J., & Rifkin, B. The development of analogical reasoning processes. Journal of Experimental Child Psychology, 1979, 27,

Footnotes

[1]The college students unfortunately do not come from the same school
population as the elementary and secondary school students. This is a com-
mon and probably insoluble problem in much developmental research, however,
caused by the failure of community school populations to remain intact after
high school.

[2]Subjects also rated the closeness of the $\underline{A}$ to $\underline{B}$ relationship to each
of the $\underline{C}$ to $\underline{D}_{Nonkeyed}$ relationships, in order to provide subjects with a con-
text in which to rate the range of relationships. These latter ratings, how-
ever, were not used in the data analyses to be reported.

[3]Subjects also rated the associative relatedness between the last term
of the analogy stem and the first term of each nonkeyed answer option. These
ratings showed the same patterns of correlations as those for association
to the keyed answer, and hence were not used in the final data analyses.

[4]Predictors were numbers of encodings, inferences, mappings, and applications, plus ratings of association and justification. In the correlations reported in the table, the 180 analogies were collapsed over the 5 semantic relations, so that each correlation was based upon 36 data points.

[5]Note that although ratings of association were collected from adults, significant correlations of these ratings with dependent variables were obtained only for children.

[6]Multiple correlations were based upon the same 36 data points upon which the simple correlations were based.

[7]The SPSS computer program used for the multiple regressions uses partial correlations as the criterion for stepwise inclusion of a variable in the regression. Other computer programs, based upon slightly different algorithms, may use part rather than partial correlations.

Table 1

Simple Correlations of Model Predictors with Response Time and Error Rate

		Response Time				Error Rate			
					Grade				
		3	6	9	C	3	6	9	C
Predictor	Model								
	A	.55**	.58**	.88**	.89**	.51**	.43**	.51**	.01
Encoding	B	.62**	.68**	.83**	.86**	.42**	.49**	.57**	.08
(No. Terms)	C	.75**	.70**	.69**	.63**	.70**	.59**	.59**	.23
	D	.83**	.84**	.73**	.69**	.66**	.66**	.65**	.17
	A	.03	.13	.64**	.73**	.10	.06	.19	-.14
Inference	B	.12	.27	.67**	.77**	.13	.17	.28	-.09
(A → B)	C	.05	.07	.39**	.49**	.22	.20	.21	.01
	D	.12	.27	.67**	.77**	.13	.17	.28	-.09
	A	.63**	.62**	.80**	.79**	.55**	.50**	.52**	.07
Mapping	B	.71**	.69**	.78**	.78**	.45**	.54**	.59**	.11
(A → C)	C	.62**	.54**	.40**	.32	.60**	.57**	.41**	.11
	D	.76**	.73**	.63**	.58**	.59**	.61**	.55**	.17
	A	.67**	.65**	.63**	.58**	.60**	.46**	.53**	.25
Application	B	.65**	.68**	.59**	.56**	.43**	.46**	.52**	.17
(C → D)	C	.56**	.46**	.12	.01	.58**	.50**	.32	.16
	D	.65**	.61	.26	.16	.50**	.46**	.44**	.33*
Justification (D → D')		.10	.15	-.07	.01	.04	.19	.22	.25
Association[a] (Stem → D$_{Keyed}$)		.50**	.45**	.08	-.02	.41**	.44**	.30	.29

[a]Lower numerical values for the association predictor indicate higher degrees of association.

*p < .05

**p < .01

37

Table 2

Multiple Correlations of Two Best Model Predictors with Response Time

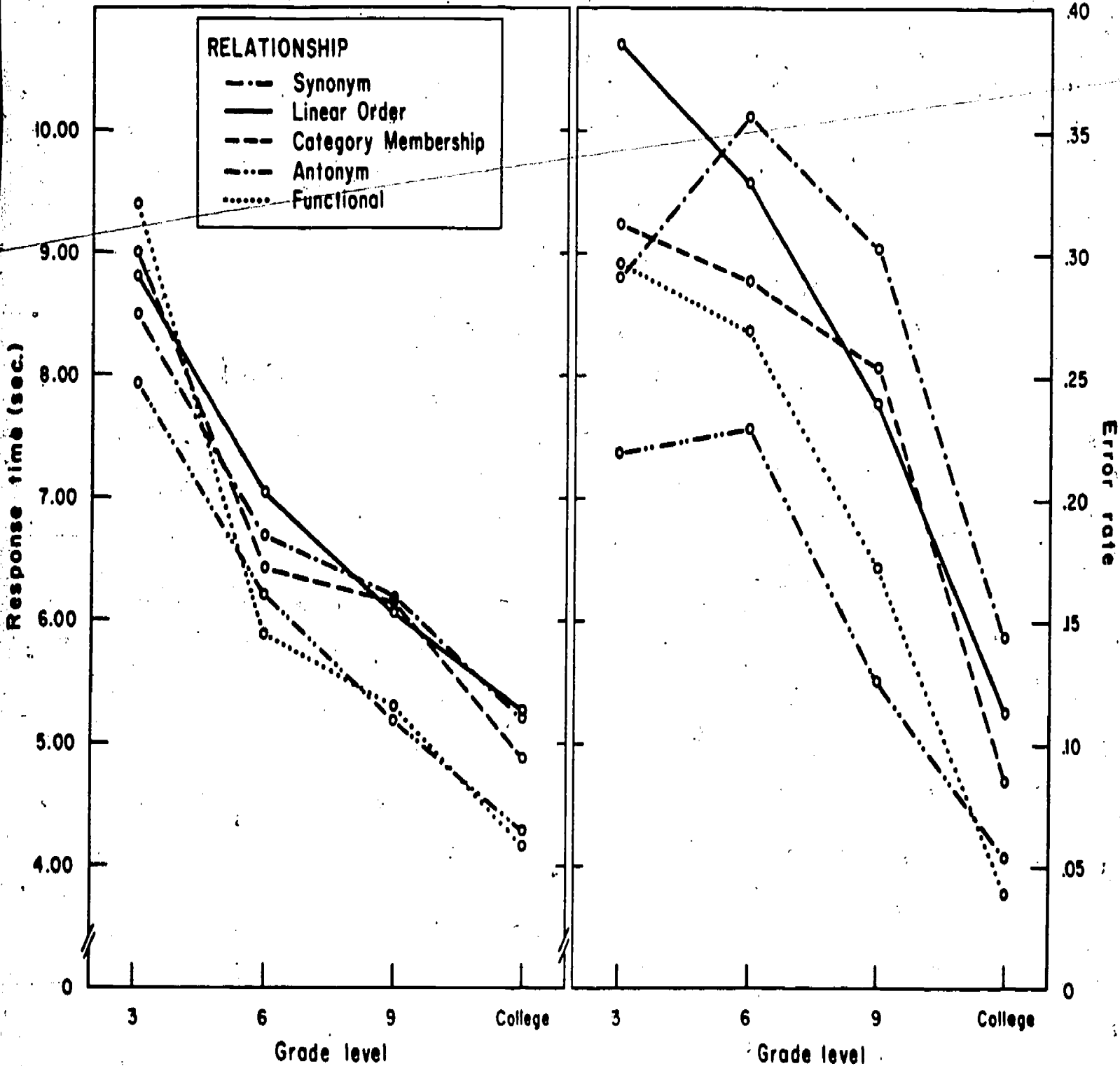| Model	First Variable		Second Variable			
	Predictor	β	Predictor	β	R	$\sqrt{r_{xx}}$
			Grade 3			.91
A	Application	.66**	Association	.47**	.82	
B	Mapping	.94**	Inference	-.41**	.79	
C	Encoding	.92**	Inference	-.36**	.82	
D	Encoding	.85**	Justification	.21*	.85	
			Grade 6			.91
A	Application	.63**	Association	.43**	.78	
B	Mapping	.43**	Application	.42**	.77	
C	Encoding	.84**	Inference	-.31*	.75	
D	Encoding	.88**	Justification	.26**	.88	
			Grade 9			.96
A	Encoding	.92**	Justification	.17*	.89	
B	Encoding	.99**	Application	-.19	.84	
C	Encoding	1.08**	Application	-.59**	.82	
D	Encoding	.56**	Inference	.47**	.85	
			College			.97
A	Encoding	.94**	Justification	.22**	.92	
B	Encoding	1.12**	Application	-.33*	.88	
C	Encoding	1.11**	Application	-.72**	.84	
D	Inference	.60**	Encoding	.47**	.89	

Note: β is standardized regression coefficient; $\sqrt{r_{xx}}$ is square root of internal
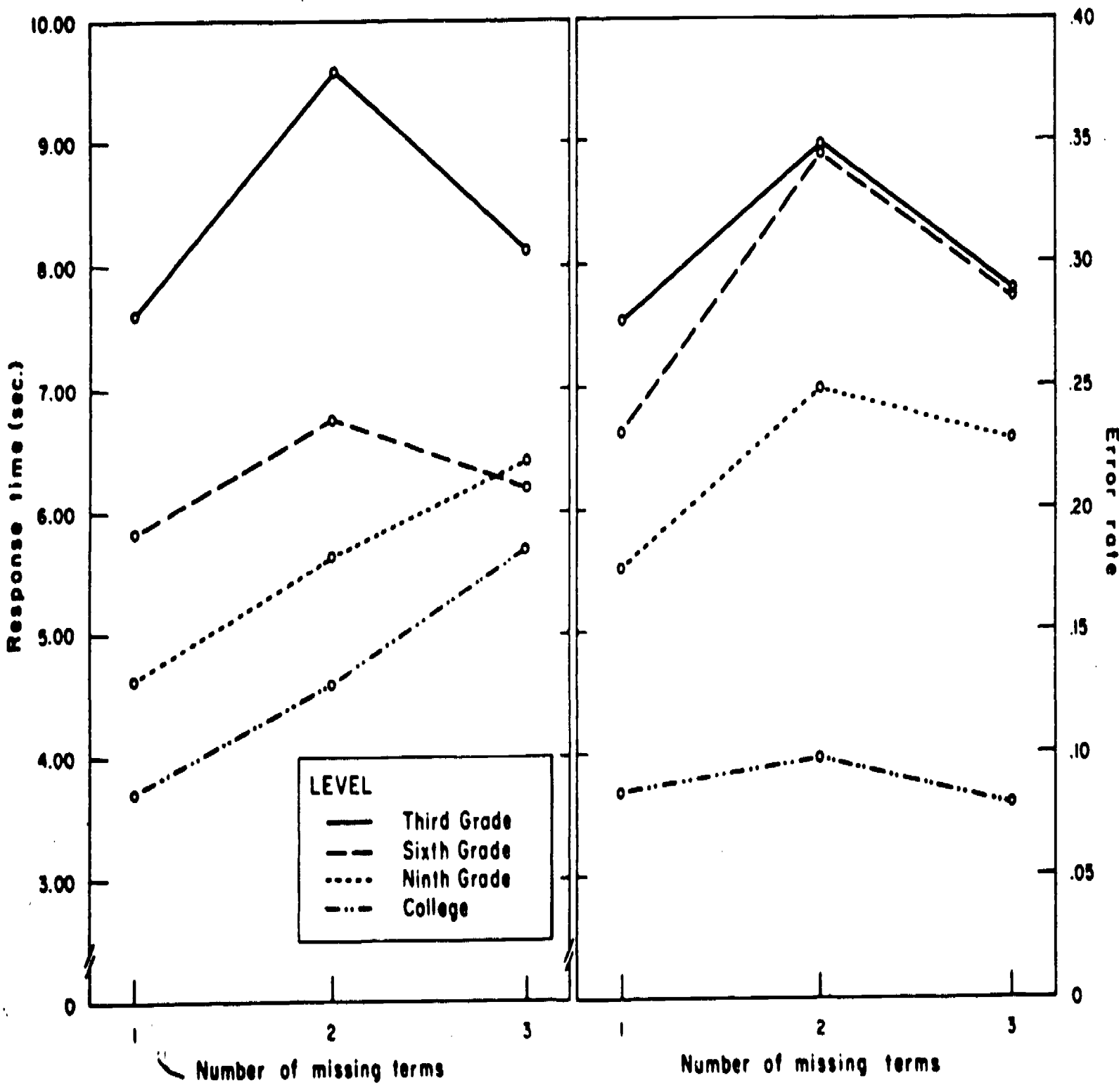consistency reliability (coefficient alpha) for latency data, the maximum possible R.

*p < .05

**p < .01

Figure Captions

1. Mean response times and error rates for each semantic relationship at each grade level.

2. Mean response times and error rates for each item format at each grade level.

RELATIONSHIP
- ─·─·─ Synonym
- ───── Linear Order
- ───── Category Membership
- ─··─··─ Antonym
- ·········· Functional

Response time (sec.)

Error rate

Grade level

Grade level

40

Response time (sec.)

10.00

9.00

8.00

7.00

6.00

5.00

4.00

3.00

0

LEVEL

Third Grade
Sixth Grade
Ninth Grade
College

Number of missing terms

1   2   3

.40

.35

.30

.25

.20

.15

.10

.05

0

Error rate

Number of missing terms

1   2   3

41

#1. Sternberg, R. J., & Rifkin, B. The development of analogical reasoning processes. June, 1978.

#2. Sternberg, R. J. Developmental patterns in the encoding and combination of logical connectives. October, 1978.

#3. Sternberg, R. J., & Nigro, G. Developmental patterns in the solution of verbal analogies. April, 1979.

#4. Sternberg, R. J. The development of human intelligence. April, 1979.

#5. Sternberg, R. J. The development of linear syllogistic reasoning. April, 1979.