Technical Report No. 279

ASSOCIATIVE ERRORS IN CHILDREN'S ANALOGICAL REASONING: A COGNITIVE PROCESS ANALYSIS

William C. Tirre

Air Force Human Resource Laboratory
Brooks Air Force Base, Texas

June 1983

Center for the Study of Reading

TECHNICAL REPORTS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
51 Gerty Drive
Champaign, Illinois 61820

BOLT BERANEK AND NEWMAN INC.
50 Moulton Street
Cambridge, Massachusetts 02238

The National Institute of Education
U.S. Department of Education
Washington, D.C. 20208

CENTER FOR THE STUDY OF READING

Technical Report No. 279

ASSOCIATIVE ERRORS IN CHILDREN'S ANALOGICAL REASONING: A COGNITIVE PROCESS ANALYSIS

William C. Tirre

Air Force Human Resource Laboratory
Brooks Air Force Base, Texas

June 1983

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02238

This article is based on the author's doctoral dissertation submitted to the University of Illinois. The research reported herein was supported in part by the National Institute of Education Contract No. NIE-C-400-76-0116. The author is grateful to Lloyd G. Humphreys, Joseph Campione, Robert Linn, and George McConkie who supervised this research as members of his dissertation committee. Thanks are also due to Peter Johnston, Brenda Lerner, Kathleen Mintert, Ann Meyer, Larry Shirey, and Peter Winograd for their help in data collection.

Abstract

Pre-adolescents often come to analogy solution via free association instead of logical reasoning, and this tendency has been related to non-adaptive learning strategies and slower intellectual growth. The purpose of this study was to investigate the cognitive processes underlying the associative response strategy in analogy solution. 112 fifth graders were administered a battery of tests designed to assess different components of analogical reasoning. The Children's Associative Responding Test (CART), a verbal analogies test which yields associative and non-associative error scores, was also administered. Factor analysis of this battery resulted in four primary factors: vocabulary, encoding and retrieval processes, discovery of semantic relations and semantic flexibility, and response evaluation. A higher order general factor was also found. Further regression analyses showed that only the mapping relations component did not significantly predict the two CART scores. Despite considerable criterion overlap, vocabulary and discovery of semantic relations were more highly related to non-associative errors, and working memory and semantic flexibility were more highly related to associative errors.

Associative Errors in Children's Analogical Reasoning: A Cognitive Process Analysis

A very common error in children's attempts to solve verbal analogies is to respond with a word strongly associated with the third term in the analogy. For instance, in "dog is to puppy as cow is to ----," many children will respond with "milk," a strong associate of "cow" but an incorrect answer. This has come to be known as the associative response phenomenon. While some researchers have argued that association is the primary component of analogy solution for all age groups (e.g., Gentile, Tedesco-Stratton, Davis, Lund, & Agunanne, 1977; Willner, 1964), the empirical evidence supports a developmental shift in strategy. For instance, Achenbach (1971) found that associative errors of this type decreased as adolescence progressed. Similarly, Sternberg and Nigro (1980) found that third and sixth grade students relied heavily upon association to solve analogies whereas ninth grade and college students relied instead upon inference.

The associative response phenomenon appears to have significance within developmental level as well. Evidence has accrued suggesting that students who make more associative errors than non-associative errors achieve less well in school as measured by grade point average (Achenbach, 1969), with this achievement gap increasing with time, as longitudinal studies have shown (Achenbach, 1971, 1975). Moreover, there is indication that level of associative responding moderates the school achievement-intelligence relationship (Achenbach, 1970a, 1970b, 1971; Tirre, Note 1). From the Tirre (Note 1) analyses it was found that associative responders were predicted to achieve less than non-associative responders of equal intelligence in reading, language arts, and mathematics. These findings corroborate earlier studies by Achenbach (1970a, 1970b, 1971).

A study by Kerner and Achenbach (1971) suggests that associative responders employ processes different from those of non-associative students when attempting to learn in school. They found that the best predictors of grades for associative responders were two rote-associative tasks: recall of categorizable items and recall of non-categorizable items. Two tasks involving reasoning, i.e., concept formation and paragraph comprehension, were not predictive at all. Precisely the opposite results were found for the non-associative students. Interestingly, the recall tasks were substantially correlated for the associative students and uncorrelated for the non-associative students, though the difference between these correlations was not quite significant. This latter finding suggests that associative students could have approached the two lists in like manners, perhaps not taking advantage of the structure in the categorizable lists. Taken together, these results imply that students who employ the associative strategy in verbal analogy problems may also fail to employ conceptual processes in other appropriate learning situations. If more were known about the cognitive nature of the associative strategy we would be in a better position to explain existing data and to make more informed hypotheses about the learning processes of associative students.

The purpose of this study was to compare and contrast the cognitive components of associative and non-associative errors on analogies as

measured by the Children's Associative Responding Test (CART) (Achenbach, 1970a). The CART consists of 68 verbal analogies, half of which include an associative foil, i.e., a distractor highly associated with the third term of the analogy. The other half of the items do not have strong associates as distractors so that association should not facilitate or impede solution. The two item types induce the student to attempt solution in different ways, i.e., in an associative manner by the associative foil items, or in a more logic-based manner by the items in which associations are less available. By examining the differences between these scores, much can be learned of the unique nature of associative responding.

The Sternberg (1977) componential framework was adopted in this study. Previous studies employing this framework but different methodologies have demonstrated its usefulness (Whitely & Barnes, 1979). In this application, the following were considered as components of analogical reasoning: semantic knowledge, working memory, encoding and retrieval, semantic flexibility, inference, mapping relations, and the response evaluation part of the application component. Each of these components could be hypothesized to be the locus of processing failure leading to the associative response. Within some of these components there could be several mechanisms or subcomponents responsible for processing failure.

<u>Semantic Knowledge Hypothesis</u>

First of all, it could be hypothesized that the associative response is due to a failure to understand the analogy terms. Inadequate semantic knowledge precludes meaningful comparison of analogy term attributes in inferring and mapping relations as well as response evaluation. In an experimental study vocabulary difficulty should be controlled; in a correlational study, it must be partialed out as an initial step.

Encoding Hypotheses

Three hypotheses concerning the role of encoding in the associative response can be entertained. The first of these is that despite adequate semantic knowledge, semantic encoding of analogy terms is too "shallow" and this leads to associative errors (Craik & Lockhart, 1971). All subsequent component operations would be adversely affected by inadequate encoding. If encoding is too "shallow," i.e., the semantic attributes of the analogy terms are not properly accessed and attended to, then we should expect subsequent operations, which are carried out upon the products of encoding, to suffer as well.

A related hypothesis is that associative responding results when the student is inflexible in thinking about the meanings of words. The primary meaning of a word might be accessed and then be too difficult to discard when it is the secondary meaning of a word that is really needed.

Yet another encoding-related hypothesis is that limited working memory adversely affects the encoding process. Smaller capacity would make the encoding process more difficult, and make attribute comparison processes in inferring and mapping relations more difficult as well. Encoding of analogy terms may have occurred without mishap. However, if the student cannot keep these attributes in consciousness, then inference and other processes will be adversely affected.

Inference Hypotheses

The fifth hypothesis is that a faltering of processing during the inference stage leads to the associative response. Consider the typical analogy form A is to B as C is to D. If a student has only a vague notion of how A and B are related in the domain (A is to B) then she/he will have a lower criterion of acceptability for a relationship between C and D in the range.

There are at least three ways in which the inference process could lead to the associative error. First of all, students may have a "conceptual style" that predisposes them to look for a particular kind of relationship at the expense of other types of relationships (Sigel, 1967). Sigel describes three different styles. The "relational-contextual" style would appear to be most at odds with analogy solution. In this case, the child groups objects together because they are functionally or thematically interdependent, e.g., horse and coach go together because the horse pulls the coach. Contrast this with the "inferential-categorical" style in which sortings are made on the basis of some inferred, shared feature. In a pilot study with 29 sixth-grade students, the Sigel (1967) Conceptual Style test was administered along with the CART. No significant correlations were found between the tests, ruling out any role of conceptual style in associative responding.

However, we may still hypothesize that faulty inferences are made when the child has inadequate knowledge of the types of semantic relations typically found in analogies. Whitely (1977) identified seven types of semantic relationships in analogies using latent partition analysis.

Instruction on these relationships can improve analogy performance (Whitely & Dawis, 1974) suggesting that the relational education or inference process is guided in some manner by knowledge of what kinds of relations are likely to be found. An obvious example of this is the problem in which a relation between "pot" and "top" is to be discovered and then applied to "ton" to complete the analogy. Knowledge that word pattern analogies are legitimate types would direct the person away from semantic comparison of attributes and towards orthographic comparison leading to the answer "not."

The third way the inference process could go awry and affect later processing is through a failure to compare and contrast semantic attributes, given that the student is aware that this is the appropriate strategy. In Sternberg's theory, inference is a matter of comparing the attribute lists of the A and B terms. For instance, "wolf" and "dog" share a number of attributes subsumed under the concept "canine." Let us suppose that on only one dimension "tameness" do they really differ. List comparison allows the inference "a wolf is like a wild dog." The hypothesis then, is that skill at such semantic processing should be highly negatively correlated with associative errors.

Mapping Hypothesis

We may also hypothesize that associative responding is the natural consequence of treating the analogy range as an isolated word pair, i.e., the relation found in the domain is never mapped onto the range. Gallagher and Wright (Note 2, 1979) argue that analogy errors result from an inadequate understanding of higher order relationships, i.e., relations between relations. They noted that symmetric explanations of analogies,

i.e., explanations which demonstrate the symmetry or balance between domain and range, are correlated highly with correct solution and increase in frequency and in sophistication as the child leaves childhood and progresses through adolescence (see also Levinson & Carpenter, 1974). As examples of symmetric explanations consider the following seventh-grade responses to Engine is to Car as Man is to Bicycle:

"Because man is a bicycle's engine."

"The first word provides power to the second."

Contrast these rule-specifications to the following fourth-grade responses which focus on the analogy range:

"A man makes a bike go."

"A man rides a bike."

In the present study, the Gallagher and Wright Written Analogical Reasoning Test (WART) was employed to determine the relation between the understanding of symmetric relations and associative responding.

Impulsiveness Hypothesis

Lastly, it was hypothesized that carelessness in evaluating alternative solutions to the analogy results in associative errors. Such carelessness could be another manifestation of an impulsive cognitive style (Kagan, Rossman, Day, Albert, & Phillips, 1964). To test this hypothesis, scores on the Matching Familiar Figures test were correlated with the CART criteria. Impulsiveness is indicated by fast, inaccurate selection of figures and reflectiveness is indicated by slow, accurate selection. A significant interaction between speed and accuracy should be noted if impulsiveness plays a role in associative responding.

Each of the six components of analogical reasoning was measured by a paper-and-pencil test. Two questions were asked of the data. The first was whether the tasks representing the components did indeed cluster in the hypothesized manner. This was determined with factor analysis. The second question concerned how the CART associative and non-associative scores are distinguished in terms of contributions by each of the components. Regression models were employed to answer this question.

Method

Subjects

The sample consisted of 127 fifth-grade students. Fifty-five children were tested in a parochial school in suburban Chicago and seventy-two children in a public school in suburban St. Louis. Nearly all the children in the Chicago area school were White, whereas about 50% of the children in the St. Louis school were Black. Deleting cases with incomplete data left 112 cases.

Procedure

Nine tests were administered to the children in their regular classrooms. From these nine tests were derived two measures of vocabulary knowledge, three measures of verbal inductive reasoning, a measure of the mapping component, and latency and error scores on the Matching Familiar Figures Test assessing impulsiveness-reflectiveness. Scores on two tests of vocabulary were obtained from school files. Two measures of different aspects of analogical reasoning were obtained from the CART: associative and non-associative errors. Testing was carried out in two one-hour

sessions, one in the morning and one in the afternoon. With the exception of the Matching Familiar Figures test, all tests were group-administered.

For all tests instructions were written to explain the task to the student, giving several examples which could be worked on individually and then together as a group. Time limits were announced for the tests.

Experimental Tasks

Measures of semantic knowledge. The Vocabulary test was designed to assess two aspects of semantic knowledge. The first aspect measured is the ability to select the meaning of a word from among several close alternatives (VOCABR). The second aspect measured is semantic flexibility (SEMFLEX), or the ability to find a second meaning of a word which is less common and which is embedded in a context more consistent with the primary meaning of the word. A sample item will make this clearer:

Fire: flames smoke water hydrant shoot.

"Flames" is closest to the primary meaning of fire. It is embedded among other words consistent with the context of flames, e.g., smoke, water, hydrant. The secondary meaning of fire is "shoot," as in to shoot or fire a gun. Instructions were to circle the first meaning found one time, and the second meaning, if found, two times. This test was inspired by a test devised by MacGinitie (1970) to measure "flexibility in word meaning."

Items for this test were in part based on research on polysemous words conducted by Mason, Knisely, and Kendall (1981). In this study the primary and secondary meanings of words were determined empirically and the effects of polysemy on reading comprehension were noted. For the present test, distractors were written to be thematically consistent with the primary meaning of the word. Interspersed among 20 items with double meanings were 10 items with only one correct answer. The test was administered in two separately timed halves of 5 minutes each.

Also measuring semantic knowledge were the Non-Literal and Literal Vocabulary scales from the Science Research Associates Primary Achievement battery. These scores were available from both schools´ files. The Non-Literal (VOCABNL) items required comprehension of word meaning in figurative and idiomatic expressions. In contrast, the Literal (VOCABL) items required comprehension of words´ most literal senses.

Measures of encoding and retrieval. Three instruments were designed to assess different aspects of encoding and retrieval. The Same or Different task (ENCRET1) presented the student with two lists of 32 word pairs. The task was to circle "Same" if the words had the same or similar meanings, and "Different" if they had clearly different meanings. These lists were presented with one minute time limits and instructions which stressed speed and accuracy. Of the 64 word pairs on the lists, only 20 pairs did not contain near-synonyms. In these pairs, the first word was followed by a high frequency associate with a distinctive meaning. All words were selected from standardized vocabulary tests designed for third through fourth graders, thereby lessening the role of vocabulary in decision time. The resulting score was correct semantic decision rate, reflecting the speed with which words could be read in, meanings accessed and represented in working memory (encoded), meanings compared, plus speed of motor response.

Two other measures of encoding and retrieval processes were embedded in an inductive reasoning task. The goal was to construct a measure of the quality of encoding that transpires when the person is reasoning inductively such as in a verbal analogy task. If words are being processed in an appropriate semantic fashion, i.e., attributes are being accessed and represented in working memory, then memory for these words should be stronger. A verbal classification test was selected because this task involves inductive reasoning and it happens to resemble a categorized word list, adaptable for use in an incidental recall task. Of particular interest here was the clustering index. Clustering of items of similar meaning or clustering by category membership would indicate that items have been organized in memory according to shared semantic features, clear evidence of semantic analysis of the stimulus words.

This test was labeled "Which Word Does Not fit?" and consisted of four parts. Part I consisted of eight verbal classification problems in which the task is to pick the one word which does not belong with the other four words in the group (time limit: 2 minutes). Part II was a surprise free recall task. Students were instructed to write down as many words from the word groups as they could remember (time limit: 3 minutes). Parts III and IV consisted of 15 verbal classification items each (time limit: 3 minutes).

The free recall task wielded two measures: a total correct recall score (ENCRET2) and a clustering score (PCCLUS). Parts III and IV of the test were used as measures of verbal inductive reasoning, one aspect of discovering semantic relations (DSR1).

Measures of discovering semantic relations. In addition to DSR1, the Word Grouping Game (DSR2) and "How Are These Words Related?" (DSR3) were included as measures of skill at discovering semantic relations.

The Word Grouping Game consisted of two sets of seven words which could be sorted into groups of varying sizes according to different shared attributes. The first word set consisted of seven living creatures and the second set, seven items of food. In separately timed sections (5 minutes each), students were instructed to write the letter of each word to be included in the group and then explain what the shared attribute was. Students began the task by working through detailed instructions with the experimenter who explained what a valid group would be in several examples.

The "How Are These Words Related" test (DSR3) consisted of two lists of word pairs representing six of the seven different semantic relations in analogies identified by Whitely (1977). Whitely's seventh relation, the word pattern, was not included. The eight relations were the following: antonyms, synonyms, functional, quantitative, conversion, class-naming, causation, and property/feature. The last two types were identified in Millman and Pauk (1969). Thus, the major types of semantic relations in analogies were represented in the test. Students were instructed to write a short sentence explaining how the two words were related (time limit: 4 minutes per list). For example, for "cat" and "kitten" a student could write that a kitten is a baby cat.

Measures of mapping relations. The Written Analogical Reasoning Test (Gallagher & Wright, 1979) was obtained from the authors as a measure of the mapping component (MAPR). The WART consists of two parts each with 10

multiple choice analogy items. Ten items have concrete type relations and ten have abstract relations. The task is to solve the analogy and then explain or justify one's choice. For this study the test was renamed the Solve and Explain test. New instructions were written to enable the experimenter to demonstrate different forms of explanation that students could use. Students worked through two examples and discussed each. Students finished well before the 12 minute limit.

Measures of response evaluation. To measure impulsiveness/ reflectiveness in response evaluation the traditional test was chosen, i.e., the Matching Familiar Figures test (Kagan et al., 1964). In this task, students are presented a target picture and six alternatives from which they are to select the one picture which matches the target identically. The six alternatives are all very similar to one another, requiring the student to carefully evalute each one.

This test was administered individually to the students. Latency to first response (MFFT) and total number of errors (MFFE) were recorded. Students were instructed to work on each item until they found the right answer.

Measures of working memory. To measure working memory capacity a digit span memory test was devised (Case, 1974). The task was presented as a game "How Many Numbers Can You Remember?" (DSPAN) and was administered to the entire class, with trial one in the morning, and trial two in the afternoon.

The experimenter read aloud seven lists of digits, starting with a 3 digit list and ending with a 9 digit list. Before the test the experimenter practiced reading the lists silently inserting the word "thousand" between digits to approximate a one second interval. As the students listened to the lists, they were instructed to raise their arms with pencil in hand to prevent any writing. After the last digit, students attempted to reproduce the number sequences on response sheets.

Results and Discussion

Test Scoring

Most of the 13 tests in the experimental battery could be objectively scored. Of these tests, only the incidental free recall task and experimental vocabulary test require further explanation.

The experimental vocabulary test yielded two scores. The regular vocabulary score (VOCABR) was the number of correct primary meanings selected from the 20 items with double meanings plus the number of correct meanings selected from the 10 items with single meanings. The flexibility score (SEMFLEX) was the number of correct second meanings selected from items in which a primary meaning was also selected. In other words, a circled secondary meaning counted toward the flexibility score only if the primary meaning was also circled.

The incidental free recall task was scored for both number of words correctly recalled and degree of clustering. The former was scored as the number of verbatim list words written on the test page, counting misspelled words but not synonyms. A ratio measure of clustering was chosen after considering the recommendations of Murphy (1979). The measure chosen was the simple percentage of words recalled in clusters (i.e., words of same category grouped together). This measure correlates .95 with the ratio of repetition but less of its variance is due to confounding variables.

Tests which required judgment on the part of the scorer included the Solve and Explain test, the "How Are These Words Related?" test and the Word Grouping Game. Detailed scoring guides were constructed for each of these tests and 25 test papers were randomly sampled from the 112 papers. These were scored by another person trained by the experimenter. Inter-scorer agreement was 93.75% for the "How Are These Words Related?" test and 92.6% for number of valid groups listed on the Word Grouping Game.

Scoring instructions for the Solve and Explain test were modified somewhat from the original WART instructions. In the original system an explanation of an analogy was scored as either symmetric or asymmetric, with no middle ground. The revised system scales response on a three point scale. Receiving full credit as symmetric responses are rule reason or successive reason explanations (Gallagher & Wright, Note 2, 1979). Receiving half credit are responses which do indicate some understanding of the analogy but fail to fully demonstrate the symmetry which exists between domain and range. The relation expressed could apply to both range and domain, but the student does not bother to demonstrate this, focusing only on the range. Receiving zero credit are the responses which fail to compare domain and range, display inversion (A:B::D:C), or appear to state an associative rule for the answer, e.g., "C and D go together." It should be noted that asymmetric justifications could be given to correctly solved analogies. This overall symmetric explanation score correlated .75 with correct analogy solution.

Inter-scorer agreement was again very high. It was 96% for the symmetric explanation category, 88.3% for the range--focusing category, and

95.9% for the asymmetric--no comparison category, the most frequent category. Association proved to be too difficult to distinguish from the asymmetric--no comparison category.

Reliability estimates for the various tests appear in Table 1. With the exception of the SRA vocabulary tests, the estimates reported are split-half correlations corrected with the Spearman-Brown formula. The Kuder-Richardson Formula 21 was used to estimate the SRA tests' reliabilities.

Insert Table 1 about here.

Factor Analysis of the Analogy Solution Components

The correlation matrix for the 14 analogy cognitive components was factored using the principal axes method. Initial estimates of the communalities were squared multiple correlations. The number of factors to retain and rotate was decided by the parallel analysis criterion (Humphreys, Ilgen, McGrath, & Montanelli, 1969). This criterion accepts as meaningful only those factors with an eigenvalue greater than the corresponding eigenvalue of a matrix of correlations among random numbers. These random eigenvalues can be estimated using a regression equation published in Montanelli and Humphreys (1976). In this case, the random data eigenvalue for factor five exceeded that of the real data factor five, so four factors were rotated. Four factor solutions were also indicated by the Kaiser-Guttman unity criterion and by the maximum likelihood chi-square test.

An oblique factor rotation was obtained using the Binormamin program. The resulting factor intercorrelations suggested a higher order general factor. Thus, the factor correlations were themselves factored, wielding a general factor.

It was decided to use the Schmid-Leiman (1957) orthogonalization procedure which allows one to represent in one matrix the loadings of observed variables upon higher order factors and upon the primary factors. Matrix elements are correlations between the variables and that part of the primary factor which has the higher order factor partialled out. The pattern matrix *Pvo* which has *v* variables as rows and *o* orthogonal factors as columns is obtained by the formula: *Pvo* = *Pvf.Af[h + f]*, in which *Pvf* is the primary factor pattern, and *Af[h + f]* is *[Pfh|Uff]*, i.e., the higher order factor patterns augmented by a diagonal matrix whose elements are the square roots of the uniquenesses of the primary factors.

Table 2 displays the Schmid-Leiman orthogonal factor pattern for the present data. The factors can be interpreted as follows. The higher order general factor is probably best regarded as general intelligence. General intelligence can be defined as that subset of procedural and declarative knowledge which is most commonly tapped by the various cognitive tasks in academic settings. The tests which have the highest loadings on this factor are those which have been traditionally used to measure intelligence: verbal reasoning (DSR1, DSR3, MAPR) and vocabulary (NLVOC, LVOC).

Insert Table 2 about here.

The first primary factor was loaded by the error and latency to first response scores from the Matching Familiar Figure test. This factor is probably best interpreted as reflectiveness-impulsiveness in evaluating alternative solutions. In this case, the factor relates well to the response evaluation component in Sternberg's theory. The role of image generation, i.e., forming an image of the ideal answer, is probably minimized in this task since the target picture is readily available. That these scores load minimally on the general factor is at least in part due to the visual/figural content of the test. The rest of the battery involves verbal content.

The second primary factor is a combination of discovering semantic relationships and semantic flexibility, and thus, corresponds nicely to the inference component in Sternberg's theory. Discovering how words are semantically related is important in the word classification task, the word grouping task, and the identifying semantic relations task. Semantic flexibility is involved in this factor as well, as indicated by the SEMFLEX loading. Another type of semantic flexibility is measured by the word grouping task, which some authors use as a measure of "semantic spontaneous flexibility" (Hakstian & Cattell, 1974). High scores on this test are the result of overcoming the cognitive set established by the previously encoded attributes and searching for new attributes upon which new groups may be formed.

The MAPR score loaded on this factor no doubt because of the involvement of discovering semantic relations, a logical prerequisite of mapping relations. A student must be able to infer relations between two concepts before she/he can reflect on higher order relationships between relations. In the Carroll (1980) re-analysis of Sternberg's data, mapping and inferring relations also loaded the same factor. In the present study, however, mapping relations could not be expected to define its own factor since it was under-represented in the battery.

Percent clustered was included in this battery because of its sensitivity to encoding semantic attributes. A high degree of clustering in recall is, in a sense, a record of success in encoding and comparing the correct semantic attributes of words. It is not surprising that this measure loads the same factor as tasks requiring the discovery of relationships.

The two standardized vocabulary tests are the primary variables loading the third factor. Loading less well is the regular vocabulary score from the experimental test designed for this study. The standardized tests required reading comprehension skills at the sentence level, whereas the new test did not, which may explain its weak loading.

The last factor is best interpreted as the encoding and retrieval processes factor. The recall task and the semantic decision rate task load on this factor, but the clustering index does not, contrary to expectations. Speed of processing can be ruled out as an interpretation since both the study and recall phases of the incidental recall task had generous time limits allowing an unspeeded work pace. Instead, what appears to be shared by these tasks are the processes of encoding the meanings of words and retrieving information from memory.

One objection might be that the semantic decision task involves retrieval from semantic memory and that the recall task involves retrieval from episodic memory. Kintsch (1977) argues that the distinction between episodic and semantic memory traces is artificial. The conception of memory in terms of feature sets applies equally to both kinds of memory. Retrieval mechanisms are highly similar too. The semantic memory retrieval model of Smith, Shoben, and Rips (1974) is closely paralleled by the episodic memory retrieval models of Atkinson and Juola (1974) and Wescourt and Atkinson (1976). Kintsch's view is supported by the present finding.

<u>Multiple Regression Analyses</u>

Multiple regression modeling is a flexible technique which, through the hierarchical inclusion method, allows the specification of the causal priority of variables, either temporally or logically determined (Cohen & Cohen, 1975). It also allows one to test hypotheses about interactions between independent variables.

It was decided to work with composite scores representing the analogy solution components instead of factor scores obtained from the analysis reported above. It was felt that the factor analysis could have glossed over subtle differences between independent variables that multiple regression might be sensitive to. As an example consider that MAPR loaded the factor with all the DSR tests, but just may explain criterion variance left unexplained by the DSR composite score.

For this analysis all variables were transformed to standard scores and various composites were formed. Semantic knowledge was represented by a vocabulary composite consisting of VOCABNL, VOCABL, and VOCABR. An encoding and retrieval processes composite was formed with ENCRET1 and ENCRET2. Discovering semantic relations was represented by a composite of DSR1, DSR2, and DSR3. Left as single scores were SEMFLEX, DSPAN, MAPR, PCCLUS, MFFT and MFFE.

Causal priority was determined by temporal sequence. That is, since encoding processes would have to operate before an inference could be made, and inferential processes in turn, would have to operate before any mapping of relations could occur, these variables were entered in that order into the equation. After MAPR was entered, MFFT and MFFE were entered, this order following the logic that response evaluation would occur after the mapping process.

Two related regression models were tested on both the CART foil errors (CARTFE) and non-foil errors (CARTNFE). Model A entered VOCAB and DSPAN on the first step in the hierarchical inclusion process. This allowed one to determine the influence each predictor had that could not be attributed to semantic knowledge and working memory capacity. Model B analyses involved using CARTNFE as the first-entered covariate in the analysis of CARTFE, and vice versa, to permit another perspective on the data. In all analyses, the final step was the inclusion of the product terms MFFE X MFFT, PCCLUS X DSPAN, ENCRET X DSPAN, PCCLUS X DSR, ENCRET X DSR.

As it turns out, CARTNFE and CARTFE overlap substantially ($\underline{r}^2 = .46$), reflecting the operation of similar cognitive processes. General intelligence is operating in both strategies, though perhaps not to the same degree. This was indicated in the loadings of -.577 for CARTFE and -.633 for CARTNFE on g, obtained from a Dwyer extension analysis (Dwyer, 1937).

Regression analyses suggest that the two strategy scores are distinguishable in terms of componential contributions, however. The Model A analysis in Table 3 shows that for CARTFE significant increments in explained variance are present for vocabulary (.302), digit span (.097), percent clustered (.044), semantic flexibility (.045), discovery of semantic relations (.043), and MFF time (.018). These increments are squared semi-partial correlations, i.e., correlations between the dependent variable and independent variable, with the influence of previously entered independent variables partialed out.

Insert Table 3 about here.

The signs of the regression weights indicate that all relationships are in the expected directions. Vocabulary and digit span account for the most variance (30% and 9.7%, respectively) with additional increments of 4.3 to 4.5 percent added by percent clustered, semantic flexibility, and discovery of semantic relations.

The picture is somewhat different for CARTNFE (see Table 4). Significant increments in explained variance are due to vocabulary (.431), digit span (.026), percent clustered (.040), discovery of semantic relations (.119), and MFF time (.013). In this case, vocabulary and discovery of semantic relations account for most of the variance (55%).

Insert Table 4 about here.

Before addressing the significance of the differences between the CART foil and non-foil scores, several findings true of both scores should be noted. First of all, the CART scores are nearly identical with respect to the components which reliably predict performance. Contributing to both scores were semantic knowledge, working memory, encoding and retrieval (as indexed by percent clustered) discovery of semantic relations and response evaluation (MFF time). Furthermore, all of the components, with the exception of mapping relations, contribute to the explained variance. It is probably the case that the mapping relations score shared too much with the semantic relations score to explain additional criterion variance. It should also be noted that an additive model is probably sufficient for these data since all product terms failed to add significantly to the explained variance. Included here is the MFF error by latency term, indicating that impulsiveness is not likely to play an important role. Perhaps though, something akin to time spent encoding stimuli and evaluating alternatives is important, given the significant MFF time semi-partial correlation.

To test the differences between the contributions made by each component to the criteria, the t-test for the difference between two correlations for a single sample was applied to the semi-partial correlations (Ferguson, 1971). These *t*-tests show that vocabulary probably plays a more important role in non-associative errors (semi-partial *r* =

-.656) than in associative errors (semi-partial r = -.549), $t(109) = 1.8$, $p < .065$. This is also true of the DSR component (semi-partial r equals -.345, -.207, respectively), $t(104) = 1.86$, $p < .065$. On the other hand, working memory capacity plays a larger role in associative errors (semi-partial r = -.311), than in non-associative errors (semi-partial, r = -.162), $t(108) = -2.04$, $p < .05$.

Another way to examine differences between the cognitive components of associative and non-associative errors is to use one error score as a covariate in the prediction of the other. Variables entered into the equation after the covariate will show increments in explained variance not attributable to the covariate.

Tables 5 and 6 display the Model B analyses for the associative and non-associative error scores respectively. With non-associative variance controlled, associative variance is explained in increments by vocabulary (.019), digit span (.051), and semantic flexibility (.021). With associative variance controlled, significant increments in non-associative variance are found for vocabulary (.115), and discovery of semantic relations (.078). These analyses corroborate the earlier findings and suggest also that semantic flexibility plays a larger role in the events leading to the associative type error.

Insert Tables 5 and 6 about here.

General Discussion

This study lends partial support to Sternberg's theory of analogical reasoning by demonstrating that encoding and retrieval processes, semantic inference processes, and something akin to response evaluation each predict analogy performance when entered into a regression equation in a theory-specified order. This study also demonstrates the importance of vocabulary, working memory capacity, and semantic flexibility in the solution of verbal analogies.

Most importantly, it was found that the associative and non-associative error types overlap considerably in underlying cognitive processes. Every component, with the exception of MAPR, explained variance in CARTFE and CARTNFE. The strengths of these contributions differed in interesting ways, however. Semantic knowledge and discovery of semantic relations appear to have stronger roles in the events leading to the non-associative error than in those leading to the associative error. On the other hand, working memory capacity and semantic flexibility appear to have greater importance in the events leading to the associative response.

These findings can be tentatively interpreted to mean that when association is not available or is avoided in analogy solution, great reliance is placed upon semantic knowledge and the ability to reason inductively with words. Together these components account for 55 percent of the variance. When association is employed in analogy solution, this could be the result of limited working memory capacity and perhaps, inflexibility in accessing the meanings of analogy terms, or some other related kind of inflexibility. Working memory and semantic flexibility account for only 9.7% and 4.5% of the variance, respectively, so much of the unique nature of the associative strategy remains to be explained.

There are several areas that need further exploration in the search for the cognitive events leading to the associative response. One interesting finding was that working memory capacity was more strongly related to associative errors. Research has shown that associative errors decrease in frequency as the child develops (Achenbach, 1971; Sternberg & Nigro, 1980), and that working memory capacity increases (Case, 1974). A potential link between working memory and associative errors could be the inference component. Previous research has shown that limited working memory adversely affects the inference process (Kotovsky & Simon, 1973; Holzman, Pellegrino, & Glaser, 1982). With only a vague idea of how A and B are related because of an inability to effectively compare attribute lists in working memory, the student may have a lower criterion of acceptability for a relationship in the range. As a consequence, the salient, associative relationship is chosen.

A second area needing further exploration is the relationship between mapping relations and associative responding. In this study mapping relations did not add significantly to the prediction of either CARTFE or CARTNFE, when entered in its theory-specified position. Simple correlations were quite strong (-.562 for CARTFE, -.529 for CARTNFE), but because a highly correlated variable (DSR) was entered first, MAPR could not add to the explained variance.

Final assessment of the mapping relations component must be postponed until additional measures of this type of reasoning can be developed.

Under-representation in the test battery probably diminished its chances of demonstrating its unique nature and role in explaining associative score variance. One possible direction for new measures is suggested by the similarity of proportional and analogical reasoning (for review see Gallagher & Mansfield, 1980). A proportion is a kind of quantitative analogy. Recognition that 1 is to 2 as 3 is to 6 implies 1 is to 3 as 2 is to 6 indicates a higher order understanding of proportionality that parallels the understanding of analogy reflected in symmetric rule reasons given as analogy answer justifications.

In conclusion, this study has pointed to several differences between associative and non-associative errors in analogical reasoning. The next step for research should be to design an experiment in which vocabulary difficulty, ease of inference, demands upon working memory capacity, and perhaps polysemy of analogy terms are varied factorially. Such experimentation should allow a more definitive assessment of the importance of these components in associative and non-associative errors.

Reference Notes

1. Tirre, W. C. Associative errors in children's analogical reasoning: A cognitive process analysis. Ph.D. Dissertation, University of Illinois, 1983.
2. Gallagher, J. M., & Wright, R. J. Children's solution of verbal analogies: Extension of Piaget's concept of reflexsive abstraction. Paper presented at the Symposium "Thinking with the Left Hand: Children's Understanding of Analogy and Metaphor," at the bienniel meetings of the Society for Research in Child Development, New Orleans, 1977.

References

Achenbach, T. M. Cue learning, associative responding, and school performance in children. Developmental Psychology, 1969, 1, 717-725.

Achenbach, T. M. The CART: A possible alternative to group IQ tests. Journal of Educational Psychology, 1970, 61, 340-348. (a)

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

Achenbach, T. M. The Children's Associative Responding Test: A two-year follow-up. Developmental Psychology, 1971, 5, 477-483.

Achenbach, T. M. A longitudinal study of relations between associative responding, IQ changes, and school performance from grades 3 to 12. Developmental Psychology, 1975, 11, 653-654.

Atkinson, R. C., & Juola, J. F. Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, R. D. Luce, & P. Suppes (Eds.), Contemporary developments in mathematical psychology, Vol. 1. San Francisco: Freeman, 1974.

Carroll, J. B. Individual differences in psychometric and experimental cognitive tasks (Report No. 163). Chapel Hill: University of North Carolina, L. L. Thurstone Psychometric Laboratory, April 1980.

Case, R. Structure and strictures: Some functional limitations on the course of cognitive growth. Cognitive Psychology, 1974, 6, 544-573.

Craik, F. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 671-684.

Cohen, J., & Cohen, P. Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, N.J.: Erlbaum, 1975.

Dwyer, P. S. The determination of the factor loadings of a given test from the known factor loadings of other tests. Psychometrika, 1937, 2, 173-178.

Ferguson, G. A. Statistical methods in psychology and education (3rd ed.). New York: McGraw-Hill, 1971.

Gallagher, J. M., & Wright, R. J. Structure of analogy items: A Piagetian interpretation. In J. Mugary (Ed.), Piaget and the helping professions (Vol. 8). Los Angeles: University of Southern California Press, 1979.

Gentile, J. R., Tedesco-Stratton, L., Davis, E., Lund, N. J., & Agunanne, B. C. Associative responding versus analogical reasoning by children. Intelligence, 1977, 1, 369-380.

Hakstian, A. R., & Cattell, R. B. The checking of primary ability structure on a broader basis of performances. British Journal of Educational Psychology, 1974, 44, 140-154.

Humphreys, L. G., Ilgen, D., McGrath, D., & Montanelli, R. Capitalization upon chance in rotation of factors. Educational and Psychological Measurement, 1969, 29, 259-271.

Kagan, J., Rossman, B. L., Day, D., Albert, J., & Phillips, W. Information processing in the child: Significance of reflective and analytic attitudes. Psychological Monographs, 1964, 78(1, Whole No. 578).

Kerner, M., & Achenbach, T. The Children's Associative Responding Test: Its relation to individual tests of recall, comprehension, and concept formation. Psychological Reports, 1971, 29, 119-125.

Kintsch, W. Memory and cognition. New York: Wiley, 1977.

Kotovsky, K., & Simon, H. A. Empirical tests of a theory of human acquisition of concepts for sequential patterns. Cognitive Psychology, 1973, 4, 399-424.

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

Millman, J., & Pauk, W. How to take tests. New York: McGraw-Hill, 1969.

Montanelli, R. G., & Humphreys, L. G. Latent roots of random data correlation matrices with squared multiple correlations on the diagonal: A Monte Carlo study. Psychometrika, 1976, 41, 341-348.

Murphy, M. D. Measurement of category clustering in free recall. In C. R. Puff (Ed.), Memory organization and structure. New York: Academic Press, 1979.

Schmid, J., & Leiman, J. The development of hierarchical factor solutions. Psychometrika, 1957, 22, 53-61.

Sigel, I. E. Sigel conceptual style test manual: Instructions and scoring guide. Detroit: Merrill-Palmer Institute, 1967.

Smith, E. E., Shoben, E. J., & Rips, L. J. Structure and process in semantic memory: A featural model for semantic decision. Psychological Review, 1974, 81, 214-241.

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

Sternberg, R. J., & Nigro, G. Developmental patterns in the solution of verbal analogies. Child Development, 1980, 51, 27-38.

Westcourt, K. T., & Atkinson, R. C. Fact retrieval processes in human memory. In W. K. Estes (Ed.), Handbook of Learning and Cognitive Processes, Vol. 4. Hillsdale, N.J.: Erlbaum, 1976.

Whitely, S. E. Relationships in analogy items: A semantic component of a psychometric task. Educational and Psychological Measurement, 1977, 37, 725-739.

Whitely, S. E., & Barnes, G. M. The implications of processing event sequences for theories of analogical reasoning. Memory and Cognition, 1979, 7, 323-331.

Whitely, S. E., & Dawis, R. V. Effect of cognitive intervention on latent ability measured from analogy items. Journal of Educational Psychology, 1974, 66, 710-717.

Willner, A. An experimental analysis of analogical reasoning. Psychological Reports, 1964, 15, 474-494.

Table 1

Reliability Coefficients

Variable	Reliability
SRA Literal Vocabulary (VOCABL)	.844
SRA Non-Literal Vocabulary (VOCABNL)	.848
Regular Vocabulary (REGVOC)	.755
Semantic Flexibility (SEMFLEX)	.637
Correct Semantic Decisions/min. (ENCRET 1)	.801
Number of Valid Groups (DSR 2)	.748
Verbal Inductive Reasoning (DSR 1)	.708
Identifying Semantic Relations (DSR 3)	.846
Symmetric Explanations (MAPR)	.890
MFF Errors (MFFE)	.369
MFF Time (MFFT)	.973
Digit Span (DSPAN)	.637
CART Associative Errors (CARTFE)	.888
Cart Non-Associative Errors (CARTNFE)	.827

Note. Reliability estimates for PCCLUS and ENCRET 2 are not available.

Table 2

Schmid - Leiman Orthogonalized Factors

	g	I	II	III	IV
VOCABL	.508	-.012	.029	.697	-.025
VOCABNL	.468	.034	-.039	.699	.039
VOCABR	.597	-.083	.222	.216	.091
SEMFLEX	.484	-.054	.394	.029	-.160
DSR 1	.632	.220	.343	.131	-.060
PCCLUS	.415	.108	.270	-.128	.069
ENCRET 2	.360	.037	-.027	-.001	.476
ENCRET 1	.506	-.065	.055	.033	.504
DSPAN	.348	-.085	.179	-.004	.117
DSR 3	.555	-.015	.338	-.059	.093
DSR 2	.538	.046	.334	-.039	.044
MAPR	.592	-.000	.367	.038	-.004
MFFT	.193	.754	-.031	-.012	.058
MFFE	-.231	-.732	-.041	-.024	.060

Table 3

Model A Analysis of CART Foil Errors

Variable	F	R^2	R^2 Change	rxy	Overall F	B	F
1 VOCAB	47.54	.302	.302	-.544	47.54	-.294	8.13
2 DSPAN	17.48	.398	.097	-.432	36.07	-.177	4.86
3 PCCLUS	8.60	.443	.044	-.371	21.39	-.078	.93
4 ENCRET	.31	.444	.002	-.331	-----	.013	.02
5 SEMFLEX	9.28	.489	.045	-.524	19.90	-.185	4.99
6 DSR	9.66	.532	.043	-.631	-----	-.225	2.94
7 MAPR	1.89	.540	.008	-.562	17.47	-.146	2.28
8 MFFT	4.27	.559	.018	-.253	14.46	-.162	2.85
9 MFFE	.40	.561	.002	.266	-----	-.046	.22
10 3 X 6	.29	.562	.001	-.049	9.60	.057	.39
11 8 X 9	.77	.565	.003	-.190	-----	-.042	.50
12 2 X 4	1.66	.572	.007	-.006	-----	.121	1.74
13 4 X 6	1.97	.581	.008	-.175	-----	-.157	1.96
14 2 X 3	.04	.581	.000	-.025	-----	.021	.04
Intercept						-.061	

Degrees of Freedom are 1,110 for Variable 1

$F(1,100)=2.75$ for

$\alpha = .10$

$F(1,100)=3.94$ for

$\alpha = .05$

$F(1,100)=6.90$ for

$\alpha = .01$

Table 4

Model A Analysis of CART Non-Foil Errors

Variable	F	R^2	R^2 Change	rxy	Overall F	B	F
1 VOCAB	83.08	.431	.431	-.656	83.08	-.418	10.45
2 DSPAN	5.26	.457	.026	-.312	45.86	-.024	.11
3 PCCLUS	8.58	.497	.040	-.366	27.70	.008	.01
4 ENCRET	2.56	.509	.012	-.391	-----	-.029	.08
5 SEMFLEX	5.58	.533	.025	-.489	32.79	-.113	2.30
6 DSR	34.82	.652	.119	-.747	-----	-.585	24.66
7 MAPR	.00	.652	.000	-.529	27.84	-.020	.66
8 MFFT	3.92	.665	.013	-.273	23.12	-.198	5.28
9 MFFE	1.94	.671	.006	.253	-----	-.113	1.66
10 3 X 6	.10	.671	.000	-.041	14.41	.031	.15
11 8 X 9	.00	.671	.000	-.137	-----	.000	.00
12 2 X 4	.79	.674	.003	-.044	-----	.086	1.08
13 4 X 6	.30	.675	.001	-.146	-----	-.054	.29
14 2 X 3	.12	.675	.000	-.073	-----	-.032	.12
Intercept						.014	

Table 5

Model B Analysis of CART Foil Errors

Variable	F	R^2	R^2 Change	rxy	Overall F	B	F
1 CART Non-Foil	94.07	.461	.461	.679	94.07	.274	6.19
2 VOCAB	3.96	.480	.019	-.549	50.28	-.179	2.63
3 DSPAN	11.79	.531	.051	-.432	40.70	-.171	4.74
4 PCCLUS	3.21	.544	.013	-.371	25.32	-.081	1.04
5 ENCRET	.02	.544	.000	-.331	-----	.019	.04
6 SEMFLEX	4.52	.565	.021	-.524	19.81	-.154	3.57
7 DSR	1.52	.571	.006	-.631	-----	-.064	.20
8 MAPR	2.09	.580	.009	-.562	17.77	-.140	2.22
9 MFFE	.73	.583	.003	.266	14.54	-.015	.03
10 MFFT	1.76	.590	.007	-.252	-----	-.108	1.26
11 4 X 7	.22	.591	.001	-.049	9.85	.049	.30
12 9 X 10	.81	.594	.003	-.190	-----	-.043	.54
13 3 X 5	1.17	.599	.005	-.006	-----	.098	1.17
14 5 X 7	1.69	.606	.007	-.174	-----	-.142	1.69
15 3 X 4	.08	.606	.000	-.025	-----	.029	.08
Intercept						-.065	

Table 6

Model B Analysis of CART Non-Foil Errors

Variable	F	R^2	R^2 Change	rxy	Overall F	B	F
1 CART Foil	94.07	.461	.461	.679	94.07	.221	6.19
2 VOCAB	29.55	.576	.115	-.656	74.02	-.353	14.19
3 DSPAN	.12	.576	.000	-.312	48.99	.015	.04
4 PCCLUS	3.19	.589	.012	-.366	31.42	.025	.13
5 ENCRET	2.26	.597	.009	-.391	-----	-.027	.11
6 SEMFLEX	1.52	.603	.006	-.489	31.75	-.072	.93
7 DSR	25.60	.681	.078	-.747	-----	-.535	21.11
8 MAPR	.20	.682	.001	-.529	27.59	-.012	.02
9 MFFE	.00	.682	.000	.253	22.81	-.103	1.44
10 MFFT	3.70	.693	.011	-.273	-----	-.162	3.63
11 4 X 7	.04	.693	.000	-.042	14.59	.018	.05
12 9 X 10	.04	.693	.000	-.137	-----	.012	.05
13 3 X 5	.34	.694	.001	-.044	-----	.059	.05
14 5 X 7	.04	.694	.000	-.196	-----	-.019	.04
15 3 X 4	.16	.695	.000	-.073	-----	-.040	.16
Intercept						.027	