Components of Understanding in Proportional Reasoning: A Fuzzy Set Representation of Developmental Progressions

Colleen F. Moore and James A. Dixon
University of Wisconsin—Madison

Beth A. Haines
DePauw University

MOORE, COLLEEN F.; DIXON, JAMES A.; and HAINES, BETH A. *Components of Understanding in Proportional Reasoning: A Fuzzy Set Representation of Developmental Progressions*. CHILD DEVELOPMENT, 1991, **62**, 441–459. The development of proportional reasoning was examined using a temperature mixture task. Each individual's task understanding was assessed by components measuring understanding of various principles of the task. Age differences were found in the mean component scores. More important, different patterns of components were found depending on whether the task was presented numerically or nonnumerically. Component patterns also depended on whether the task was presented such that subjects predicted the outcome of combining 2 containers of water at different temperatures (prediction task) or such that subjects inferred 1 of the 2 initial temperatures given the final temperature (reverse task). The results show the importance of distinguishing between intuitive knowledge and formal computational knowledge of proportional concepts and provide a new perspective on how intuitive and computational knowledge are related during development. Finally, the results also led to a new conceptualization of developmental levels as categories with fuzzy boundaries. Under this conceptualization, individuals can have different degrees of membership in "fuzzy developmental levels." This new concept preserves individual differences but also describes the sequence of development.

The problem of individual differences has been ever-present and bothersome in research on cognitive development (Cronbach, 1957; Kessen, 1960; Langer, 1970; Wohlwill, 1973), and the problem continues to receive attention (Kerkman & Wright, 1988; Siegler, 1987; Wilkening, 1988). In the present research, we provide a new approach to two related problems: How can we describe individual differences in development and, simultaneously, how can we describe the structure of development? First, we present a new model for proportional reasoning tasks, the "components of understanding" model, which is oriented specifically toward assessment of the individual's understanding. Second, we outline a new method for representing developmental progressions using Zadeh's (1965) fuzzy set concept. Using analyses of the individual profiles of component scores, we use the fuzzy set concept to show that individual differences and developmental progressions can be described simultaneously.

The components of understanding approach is an extension of prior approaches to characterizing individual differences in problem solving (Norman & Schemmer, 1977; Reed & Evans, 1987; Surber, 1980, 1984). Using primarily the ordinal features of each individual's data pattern, the approach provides measures of the subject's understanding of the principles by which the variables in a task are related. The first goal of our research was to characterize the development of proportional reasoning in terms of components, which are given below. In our experiment, subjects made judgments about the temperature of a container of water produced by combining two quantities of water differing in temperature (see

This research was supported in part by a grant to the first author from the Graduate School, University of Wisconsin—Madison. The authors are grateful to Steven M. Gzesh and Jane Rasico for assistance with data collection and analyses, and to Valerie Allen Ahl for comments on the manuscript. Address correspondence to Colleen F. Moore, Psychology Department, 1202 W. Johnson Street, University of Wisconsin, Madison, WI 53706.

[*Child Development*, 1991, **62**, 441–459. 0009-3920/91/6203-0013$01.00]

Noelting, 1980a, 1980b, and Strauss & Stavy, 1982, for similar tasks). Based on research on other proportional reasoning tasks (Ferretti, Butterfield, Cahn, & Kerkman, 1985; Reed & Evans, 1987; Siegler, 1976, 1981; Surber & Gzesh, 1984; Wilkening, 1981), there should be developmental differences in acquisition of the components. We expected that children would first understand the relations between the temperature of the added water and the final temperature (represented by components called main effect and monotonicity) and only later understand the interaction of added quantity with added temperature (represented by the crossover component).

A review by Surber and Haines (1987) showed that previous literature on proportional reasoning failed to distinguish between tasks that involve estimation or intuition and tasks that involve explicit calculation. Thus, a second goal of our research was to examine the development of the components of understanding when the task is presented in a way such that it requires estimation (i.e., no numerical measures of the variables are given) versus when the task is presented with numbers and subjects are encouraged to compute answers explicitly. We expected that components would be understood initially at an intuitive level, as shown on an estimation task, and only later would understanding of the same components be expressed computationally (Hammond, Hamm, Grassia, & Pearson, 1987; Surber & Haines, 1987). For example, a person knows by experience that the higher the temperature of the added water, the higher the final temperature will be. However, translating such knowledge into a computational scheme is a complex task.

A third goal of the present study was to examine the development of reversible operations in the context of the temperature task. In the prediction task, each subject was asked to predict the final temperature, given the temperature and quantity of one beaker (referred to as the "standard") which always contained water of constant quantity and temperature, and given the temperature and quantity of a second beaker (called the "added" water). Second, in the reverse task, each subject was asked to infer the *added* temperature, given the other values. The reverse task provides a measure of the completeness and flexibility of a person's understanding of the components. We expected to find partial knowledge (Wilkinson & Haines, 1987) such that homologous components of understanding would be more advanced in the prediction task as compared to the reverse task.

A fourth goal was to introduce and explore the utility of Zadeh's (1965) fuzzy set concept for describing both developmental levels and individual differences. Many developmental studies hypothesize that subjects progress through a sequence of distinct developmental levels of understanding for a specific domain (Hook & Cook, 1979; Kermoian & Campos, 1988; Nicholls & Miller, 1984; Noelting, 1980a, 1980b; Selman & Byrne, 1974; Siegler, 1981; Uzgiris & Hunt, 1975). Individuals are often classified discretely as members or nonmembers of developmental levels in spite of the fact that there are individual differences in the degree to which performance conforms to the specified developmental level. In the fuzzy set approach, each person has a degree of membership in a set, providing a quantitative representation of individual differences within developmental levels. We use the term "fuzzy developmental level" to denote a categorization of subjects according to their degree of membership or goodness of fit to a prototype response pattern. Thus, a fuzzy developmental level is a category with imprecise boundaries (Oden, 1977; Smith & Medin, 1981). We hypothesized a set of prototypes that describe distinct profiles or structures of component scores for the present task. Rather than examining only age group differences, the fuzzy set approach provides a way of grouping individuals with similar profiles of component scores. Thus the approach has the potential to reveal structural developmental changes in a given task domain without eliminating individual differences.

Components of Understanding in the Temperature Task

The components are easiest to explain with reference to the specific design of this experiment. The correct equation for solving the temperature prediction task is: $T_F = (Q_1T_1 + Q_2T_2)/(Q_1 + Q_2)$, where T_F is the final temperature, the Q's are quantities to be combined, and the T's are the temperatures to be combined. In both the prediction and reverse tasks, the quantity and temperature of the standard container were fixed at 3 units of 40°. Across trials of the prediction task, the temperature and quantity of the added water were varied in a 5 × 5 factorial design. The reverse task was designed analogously. Figure 1 presents the designs and correct answers for the two tasks of our experiment.

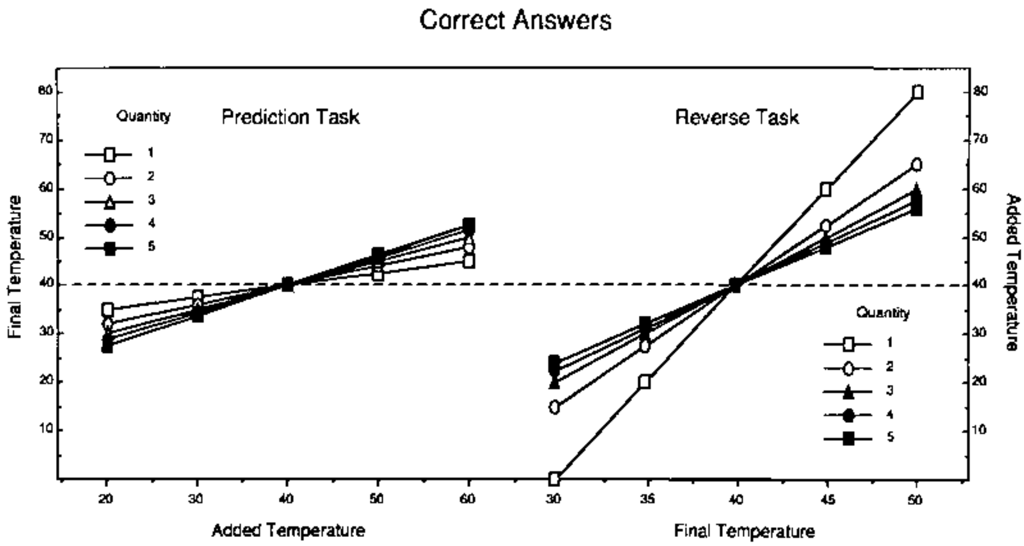

FIG. 1.—Correct answers for the prediction and reverse task plotted as a function of added and final temperature, respectively. Each curve represents a different quantity.

In research using acid mixtures, Reed and Evans (1987) scored the degree to which each individual's responses indicated an understanding of three principles of mixture tasks. We modified Reed and Evans's task components to apply to temperature mixture, and included three other components. Thus, we scored six components of understanding that measure degree of mastery of different properties of temperature mixture. Some of our components are closely related, measuring understanding at a coarse- versus a fine-grained level. The initial level of understanding a component may be measured best by the coarse-grained variable, but developmental refinements may require measurement with a finer-grained variable.

Main effect component.—The main effect component is a coarse-grained measure of understanding of the principle that the higher the added temperature, the higher the final temperature. To score main effect in the prediction task, a person was given 1 point for each correct ordering of the judgments based on the *extreme* added temperatures (20° and 60°) at each added quantity. One point was subtracted for each incorrect ordering, and nothing was done for equal judgments. The maximum and minimum scores were +5 and −5. The reverse task was scored analogously.

Monotonicity.—Monotonicity provides a more fine-grained index of the understanding that the higher the added temperature, the higher the final temperature. In the prediction task, 1 point was added to the monotonicity score if the judged temperature based on each added temperature was higher than the judged temperature for the added temperature immediately below it, holding quantity constant. One point was subtracted for each incorrect ordering of adjacent temperatures. The maximum and minimum scores were +20 and −20. To score highly on monotonicity, a person must consistently judge final temperature so that the order of responses is correct not only at the extremes of added temperature but also between each pair of temperatures. It is not completely independent of main effect. In the reverse task, monotonicity was scored analogously.[1]

Above-and-below.—This component indexes the understanding that when two temperatures combine, the resulting tempera-

[1] Reed and Evans (1987) used the term monotonicity to describe the relation between stimulus quantity (as a proportion of total quantity) and resulting acid concentration. The temperature and acid tasks are isomorphic if added and final temperature are analogous to initial and final acid concentration, and added quantity is analogous to the manipulated proportions. Thus, the monotonicity principle of Reed and Evans is *not* analogous to either main effect or monotonicity. It is closest to the component we call crossover. Reed and Evans did not measure any variables similar to main effect, monotonicity, or above-and-below. Their other two components were range and linearity.

ture is always in the direction of the added temperature. For example, if the added temperature were warmer than the standard, then the final temperature would be above that of the standard. The correct answers in Figure 1 show that in both the prediction and reverse tasks, 10 of the responses in each condition should be on each side of the standard. The score on above-and-below ranged from 0 to 20; 1 point was added to the score for each judgment in the appropriate region of the thermometer. The five given temperatures that were equal to the standard were ignored.

Range.—This component is directly analogous to Reed and Evans's (1987) range component. It is a finer-grained measure than above-and-below in that it requires the understanding that combining two temperatures results in a final temperature that is between them (i.e., within the range of the two given temperatures). In the prediction task, the range score was calculated by counting the number of responses that fell between the added temperature and the standard temperature. In the reverse task, the number of responses that fell between the given final temperature and negative infinity (for final temperatures below the standard), or between the given final temperature and positive infinity (for final temperatures above the standard) were counted. The five values of the added or final temperature that were equal to the standard were ignored. The minimum and maximum scores were 0 and 20. The range component requires that the person's responses be calibrated not only with respect to the standard but also with respect to the added temperature in the prediction task, or the final temperature on the reverse task.

Crossover.—The crossover component is essentially an ordinal index of the understanding of the interaction of quantity and temperature. Crossover was computed by checking the ordering of the responses to the two extreme *quantities* at each added temperature (for the prediction task) or at each final temperature (for the reverse task). One point was added to the crossover score for each correct ordering, and 1 point was subtracted for each incorrect ordering. The scores ranged from 4 to -4. The crossover component requires understanding the direction of change the added temperature causes, *and* that increasing the quantity of added water increases the change. For example, adding 5 cups of hot water would result in warmer water than adding 1 cup of hot water to the same standard.

Linearity.—The linearity component is also taken from Reed and Evans (1987). This component indexes the degree to which the person understands that equal changes in temperature are equal in effect regardless of their position on the temperature scale. For example, the change in added temperature from 40° to 50° has as much effect as the change from 50° to 60° for a given quantity. Linearity was computed by taking the variance of the differences of the judgments between successive given temperatures for each quantity. The linearity score increases with deviations from linearity, with an upper limit of positive infinity and a lower limit of zero.[2]

Several aspects of our approach are worthy of note. First, in contrast to much developmental research, we have not set cutoffs for "passing" or "failing" a given component. The components are graded variables that measure the degree of understanding. The vast majority of developmental studies that have examined the order of acquisition of aspects of a concept have categorized individuals into discrete levels of development; examples include the object concept or search literature (Jackson, Campos, & Fischer, 1978; Kermoian & Campos, 1988; Uzgiris & Hunt, 1975), conservation (Brainerd, 1973, 1977), proportional reasoning (Noelting, 1980a, 1980b; Siegler, 1976, 1981), and the functional measurement approach in which individuals are classified as using particular information-integration rules if certain criteria are met (Anderson & Butzin, 1978; Kun, Parsons, & Ruble, 1974; Surber, 1984). The purpose of such discrete classifications is to capture the developmental ordering of skills that are likely to be obscured by age group means. However, a number of methodological problems in determining the developmental ordering of skills from either discrete classifications or dichotomized variables have been discussed elsewhere (Brainerd, 1977; Coombs, 1964;

[2] Reed and Evans (1987) assessed linearity by using Pearson's r. Pearson's r is insensitive to deviations from linearity as long as there is a strictly monotonic relation between two variables. Insensitivity to nonlinearity can be demonstrated by calculating the r between a set of integers and their squares, square roots, or logs. In contrast, the variance of successive differences is sensitive to deviations from linearity.

Flavell, 1971; Wohlwill, 1973). In addition, once individuals are categorized into developmental levels on a task, individual differences *within* category are often ignored. The use of discrete categories of developmental levels can lead researchers inadvertently to consider development to be a more saltatory process than it actually is. We are hopeful that the fuzzy set approach will lead to a more thorough description of development than that yielded by only age group analyses or by discrete categorization of levels of development. The fuzzy set approach allows discrete categorization into developmental levels, examination of individual differences within developmental levels, and the description of developmental sequences that involve multiple paths.

Method

Subjects

Two hundred subjects from four grade levels participated: 42 second graders (mean age = 8.25 years, SD = 4.75 months), 40 fifth graders (mean age = 11.22 years, SD = 5.02 months), 48 eighth graders (mean age = 13.65 years, SD = 7.02 months), and 70 college students (mean age = 19.73 years, SD = 3.1 years). Second and fifth graders received small toys or stickers to thank them for participating. College students received extra credit points. In addition to those who completed the experiment, one eighth grader and two college students were eliminated because they were unavailable for the second session.

Approximately one-half of each age group was randomly assigned to each condition of the experiment (*N*'s = 22, 20, 23, 32 in the intuitive condition for second grade through college). Within age group and condition, approximately one-half were randomly assigned to the prediction task first (intuitive condition *N*'s = 11, 10, 12, 15, and computational condition *N*'s = 10, 10, 13, 20 for second grade through college) or the reverse task first. The proportion of males and females was approximately equal for each age group and across conditions.

Materials and Design

The stimuli in the computational and intuitive conditions were identical except that numbers were added to the display and used verbally in the computational condition, whereas qualitative terms described the temperatures in the intuitive condition. The stimuli were two felt-board beakers, approximately 12 × 16 inches in size. Blue felt represented different quantities of water that could be attached to "fill" the beakers. One of the beakers had a handle and spout that the experimenter used to simulate pouring. The schematic thermometers stood 16 inches high. The ends of the thermometers were labeled with a drawing of a fire and a snowman to symbolize very hot and very cold, respectively. A movable marker was used to delineate different temperatures. In the computational condition, the thermometers were labeled from 0 to 80 in 5° increments. In the intuitive condition, the thermometers were not graduated. Two thermometers were used to represent the two temperatures that were being combined.

The designs for the prediction and reverse tasks were 5 (added or final temperature) × 5 (quantity of added water) factorial, and are presented in Figure 1 with the correct answers. In the computational condition, temperature values were given numerically on the thermometers and verbally by the experimenter, and the numbers 1 through 5 were affixed to arbitrary equal-interval units of the felt representing the water. For both tasks, the values of the standard temperature and quantity were fixed at the midpoint of the thermometer (40° or "medium") and at approximately the one-third full point (3 units) of the beaker, respectively.

Procedure

Subjects were tested in two sessions, one for each task. The 25 stimuli for each task were presented in one of six random orders approximately counterbalanced across age and condition. The standard beaker was designated as the subject's, and always contained 3 units of water at 40°, or "medium." Subjects in the computational condition were also given paper and pencil and the experimenter said, "Try to use math to figure out the temperature. You can use the paper and pencil to help you." In the intuitive condition, neither the quantities nor the temperatures were labeled numerically. Instead, the temperatures of water were described as very hot (60°), hot (50°), medium (40°), cold (30°), or very cold (20°), with the marker on the thermometer without numbers adjusted to the same position as in the numerical task. To designate the quantity in the intuitive condition, the experimenter said, "this much water" and put it on the felt-board.

For the prediction task, the experimenter's beaker varied in the 5 × 5 design.

The subject was reminded that his or her beaker always started with the same amount of water (or 3 units) of the same temperature (medium, or 40°), but that the experimenter's water changed on each trial. The experimenter then pretended to pour her water into the subject's beaker and moved the felt strip onto the subject's beaker. The subject was asked to adjust his or her thermometer to show what the temperature of all the water together would be. In the reverse task, the experimenter first adjusted the subject's thermometer to show the final temperature after adding a certain quantity of water from the experimenter's beaker. The subject was asked to figure out the temperature of the experimenter's water prior to its combination with the subject's, and to adjust the experimenter's thermometer accordingly.

Subjects received five practice trials at the beginning of each task. Following the experimental trials, subjects were asked to explain how they figured out their answers, and were given a short arithmetic test.

Results

Age Group Analyses: Developmental Functions for Components

Figure 2 presents the mean component scores for each grade, condition, and task combination. To facilitate comparison, we transformed all the components to a percentage scale. A score of 50 represents chance performance on all the components in Figure 2 except range and linearity. Chance performance for the range component was 15.0 on the prediction task and 42.2 on the reverse task. Chance on the linearity component was 58.75.[3] The standard error bars facilitate comparisons of the differences among the means and comparisons with chance performance levels. The curves in Figure 2 are cross-sectional developmental functions for each of the components, allowing a comparison to be made of the effects of the experimental conditions on the developmental functions (see Wohlwill, 1973). Figure 2 shows that, as hypothesized, the curves for the main effect component reach a developmental asymptote quite rapidly in all four panels. Also, as expected, the crossover component is below the curves for the other components that have a chance value of 50. The above-and-below and monotonicity components appear to develop approximately in synchrony, but before crossover. The range and linearity components cannot be compared directly to the other components because their chance performance values differ, but the range component means are near chance for three age groups in the reverse task and for the second graders in the prediction task.

Intuitive vs. computational conditions.—We hypothesized that components would be acquired earlier in the intuitive version of the task than in the computational version. For the prediction task, in the top two panels of Figure 2 the mean component scores for the computational condition are generally below the means of the intuitive condition. In the reverse task (bottom panels of Fig. 2), the main effect, monotonicity, and above-and-below component means are lower in the computational condition than in the intuitive condition. The range and crossover scores are close to chance for the second through eighth graders in both the intuitive and computational conditions of the reverse task and do not show a clear pattern of differences due to condition. These effects of condition were confirmed by ANOVAs of each grade separately in a 2 (intuitive vs. computational condition) × 2 (prediction vs. reverse task) × 6 (component) design that showed significant effects of condition at all grade levels (F's = 7.50, 10.52, 5.27, 6.73, df's = 1/40, 1/38, 1/46, 1/68, for second through college, respectively, p's $<$ .05).[4]

[3] Because the chance performance level differs for two of the components, the scales for the components are not completely comparable. Calculation of chance performance for range and linearity was complex. For the range component, chance was calculated assuming a response range of 0° to 100°. For the prediction task, the average interval between the added temperatures and the standard was taken as a proportion of the total response range. This gave the average probability of a random response falling between the added and standard temperatures. Multiplying the average probability by the number of judgments yielded chance performance. For the reverse task, chance for the range component was calculated similarly. Chance for the linearity component was calculated similarly, also assuming a 0° to 100° temperature scale. The transformation of linearity to a 0 to 100 percentage scale was $(-1)\ (\text{linearity})^{1/2} + 100$. Because the original linearity scores were variances, the square root was used in the transformation. The other transformations to percentages were linear.

[4] A 4 (grade) × 2 (condition) × 2 (task) × 6 (component) repeated-measures MANOVA showed significant ($p < .05$) effects of grade, $F(3,192) = 74.92$, condition, $F\ (1,192) = 24.10$,

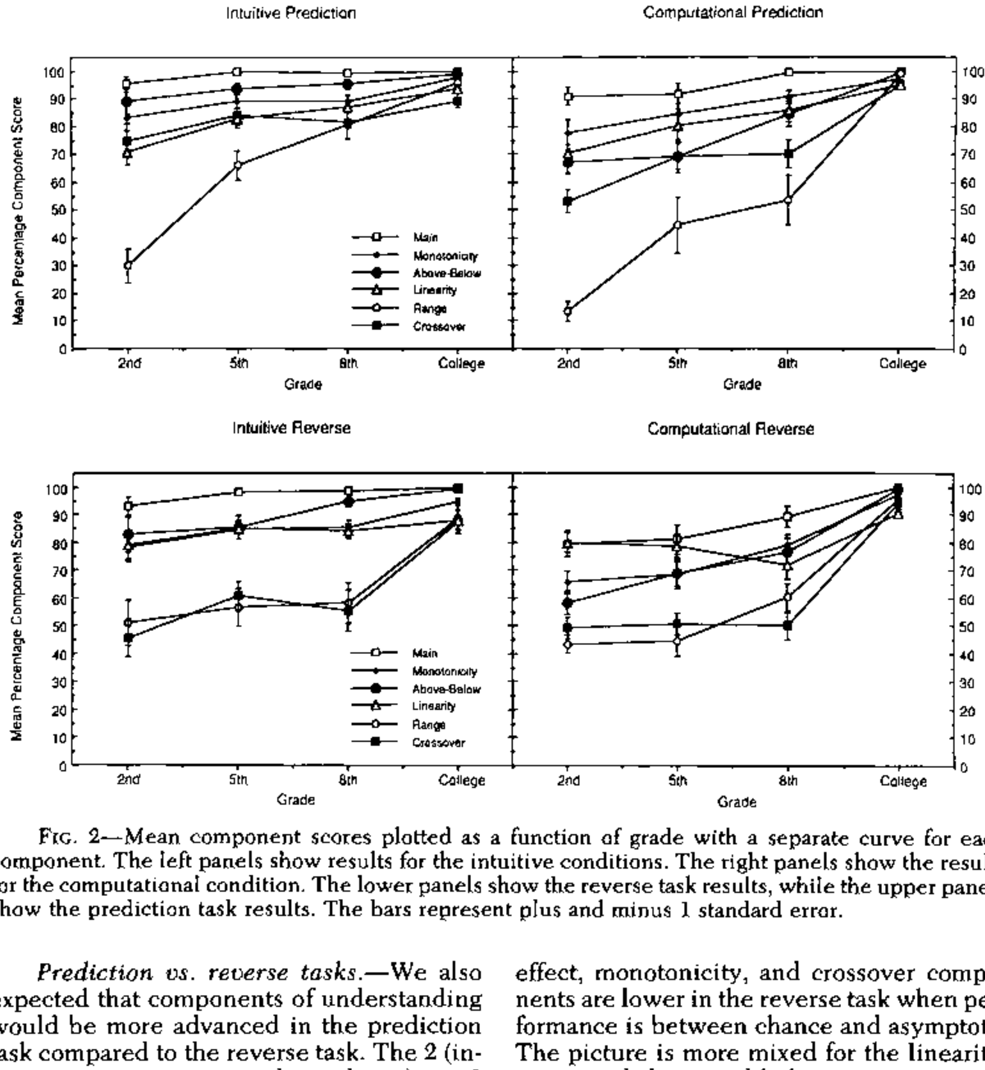

FIG. 2—Mean component scores plotted as a function of grade with a separate curve for each component. The left panels show results for the intuitive conditions. The right panels show the results for the computational condition. The lower panels show the reverse task results, while the upper panels show the prediction task results. The bars represent plus and minus 1 standard error.

Prediction vs. reverse tasks.—We also expected that components of understanding would be more advanced in the prediction task compared to the reverse task. The 2 (intuitive vs. computational condition) × 2 (prediction vs. reverse task) × 6 (component) analyses of variance of each grade level showed significant interactions of task × component for all grades (F's = 12.89, 3.83, 3.57, 3.97, df's = 5/200, 5/190, 5/230, 5/340, for second grade through college, respectively, p's $< .05$), providing evidence that the components were influenced differentially by the prediction and reverse tasks. For the computational condition, the main effect, monotonicity, and crossover components are lower in the reverse task when performance is between chance and asymptote. The picture is more mixed for the linearity, range, and above-and-below components. In sum, the expected differences between the prediction and reverse tasks were more clear for the intuitive condition than for the computational condition.

Organization of Components: Fuzzy Set Analysis of Prediction Task

The age group differences in Figure 2 could mask important individual patterns. The goal of the fuzzy set approach is to

grade × condition, $F(3,192) = 6.18$, task, $F(1,192) = 49.60$, grade × task, $F(3,192) = 8.43$, component, $F(5,188) = 183.58$, grade × component, $F(15,519) = 13.42$, condition × component, $F(5,188) = 8.14$, grade × condition × component, $F(15,519) = 1.82$, task × component, $F(5,188) = 14.85$, grade × task × component, $F(15,519) = 6.62$, condition × task × component, $F(5,188) = 3.68$, and grade × condition × task × component, $F(15,519) = 2.20$. The repeated-measures univariate ANOVAs reported in the text were conducted as simple effects tests to explore the significant interactions of grade with condition, task, and component. The p values reported in the text are all $< .05$ unless noted otherwise, and are based on Greenhouse-Geisser adjustments in the case of a repeated-measures ANOVA.

represent the individual patterns and their developmental ordering. The approach requires that the investigator define the prototype for each fuzzy developmental level. For the prediction task, we defined six prototype component patterns: Everything-Right (all component scores perfect), Everything-but-Crossover (all component scores perfect except crossover, which was set at the chance value), Everything-but-Range (all component scores perfect except range, which was set at the chance value), Everything-but-Crossover-and-Range (all component scores perfect except range and crossover), Adding (all component scores perfect except above-below, range, and crossover), and Everything-Wrong (all component scores at chance). These prototypes define a perfect member of each of six hypothesized fuzzy developmental levels.

Many possible prototypes could be chosen to represent possible developmental levels. The Everything-Wrong and Everything-Right prototypes are the obvious beginning and end points of development. The Adding prototype was included because Strauss and Stavy (1982) reported that some subjects literally added numerical temperatures. The remaining three prototypes involving differences in range and crossover were developed partly on the basis of previous findings. First, Reed and Evans (1987) found that in an acid mixture task with college students, the range and crossover components contributed independently as predictors of accuracy. Second, the age group means in Figure 2 suggest that the range and crossover components develop relatively late. Finally, following Reed and Evans (1987), examination of the frequency distributions of the components showed that both range and crossover had multipeaked distributions indicative of the presence of individual differences. Three of the prototypes allow a test of the hypothesis that range and crossover are acquired in a single developmental sequence. If there is a single sequence, then subjects should be found in either the Everything-but-Crossover or Everything-but-Range fuzzy levels, but not both.

The fuzzy set approach requires a membership function. We calculated the Euclidean distance of each individual component profile to each of the six prototype patterns, after standardizing the variables by dividing by the standard deviation. The smaller the distance to the prototype, the higher the membershp in that fuzzy developmental level. However, developmental researchers want to know the level in which an individual fits best. Using the distance measures, individuals were grouped into that fuzzy level for which distance to the prototype was smallest. The reader is reminded, however, that membership in the fuzzy levels is a continuous function.[5]

[5] Each individual has a profile consisting of the six component scores, and each prototype consists of a profile of six component scores. The Euclidean distance between the individual's profile and a prototype is:

$$d = \left[\sum_{i=1}^{6} (c_i - p_i)^2 \right]^{1/2},$$

where the c_i represent the six component scores of the individual, and the p_i represent the six component scores for the prototype. The distances were calculated on the original component scores divided by their standard deviations, not the transformed scores. We chose Euclidean distance as a measure of similarity between the subject's profile and the prototypes because it is sensitive to both the shape and elevation of a profile. Distance measures are widely used in other methods requiring similarity measures, such as clustering and multidimensional scaling, in spite of the fact that distance measures can be influenced by scale transformations. Dividing each variable by its standard deviation has the effect of changing the lengths of the axes of the multivariate space. Also, if variables that are highly correlated are used, the effect on Euclidean distance is to give higher weight to those variables. Because some of our components form hierarchical pairs (main and monotonicity, above-and-below and range), we examined the component correlation matrix. The highest correlation between any pair of components was .84 for monotonicity and above-and-below on the reverse task, not a hierarchical pair. The correlations for the hierarchical pairs were all below .8, and three of the four r's were below .7. Thus, the variables are not colinear. The hierarchical structure of the paired components restricts the areas of the multidimensional component space in which subjects can appear. For above-and-below and range, it is impossible for a person's range score to exceed the above-and-below score. The relation between main effect and monotonicity is similar in that each main effect component score sets a minimum value for monotonicity. For example, a main effect component score that is perfect limits monotonicity to be at least 20% of the maximum score. One implication of using

We hypothesized that there would be a difference between the intuitive and computational conditions in the distribution of subjects across the six fuzzy developmental levels, especially for the Adding and Everything-but-Range prototypes. Because a component profile approximating the Adding prototype is unlikely to be produced by any process other than numerical addition, subjects in the computational condition should be more likely than those in the intuitive condition to be closest to that prototype. In contrast, a profile similar to the Everything-but-Range prototype seems most likely to be produced by estimation in which temperature and quantity are used appropriately, but the subjective values of temperature are more extreme than the given values. It is difficult to devise a computational approach that would generate the Everything-but-Range pattern. However, if subjects in the computational condition resort to estimation, they could fall into the Everything-but-Range fuzzy level. Component patterns fitting the other prototypes might be produced by either estimation or explicit calculation. The Everything-but-Crossover pattern can be generated by either computation of an unweighted average or by an estimate in which both temperatures are used but quantity is ignored. The Everything-Right pattern could be produced by literal computation of a weighted average or as an estimate in which both temperature and quantity are used appropriately. Also, based on the hypothesis that performance is more advanced in the intuitive condition, especially for the younger subjects, there should be more subjects in the computational condition than in the intuitive condition who are closest to the Everything-Wrong prototype.

The left-hand columns of Table 1 present the numbers of subjects in each condition of the prediction task falling closest to each prototype. Consistent with the view that discrete categorization neglects individual differences, some individuals were equidistant from and closest to two prototypes: (*a*) Everything-Right and Everything-but-Crossover, and (*b*) Everything-but-Range and Everything-but-Range-and-Crossover. A chi-square test of independence showed a significant difference in the distribution of subjects across levels for the two conditions, $\chi^2(6) = 36.51$, $N = 200$, $p < .01$. Component profiles closest to the Adding and Everything-Wrong prototypes were infrequent in the intuitive condition compared to the computational condition, and the Everything-but-Range pattern was more common in the intuitive condition. In the intuitive condition, subjects were found in both Everything-but-Range and Everything-but-Crossover, contrary to the single developmental sequence hypothesis. Thus, the fuzzy level classification suggests that there are two developmental pathways to mature intuitive performance on this task: the range component can be acquired before the crossover component, or vice versa. In the computational condition, there were very few subjects in either Everything-but-Range or Everything-but-Crossover, implying that these prototypes are not important paths for the computational condition. This might be expected if the subjects are probabilistically selecting different computational strategies, as proposed by Siegler (1988). If subjects in the computational condition had relied largely on estimation, their distribution of memberships would have been the same as that of the intuitive condition.

Figure 3 presents the mean judgments for the predicted task for the four groups with the largest numbers of subjects. Each group has a distinct pattern of judgments that is consistent with its prototype definition, showing the success of the method in identifying groups with different response patterns. A 4 (group) × 5 (added temperature) × 5 (quantity) analysis of variance of each condition showed significant group × added temperature × quantity interactions, $F(48,1264) = 5.34$, $F(48,1312) = 2.95$ for intuitive and computational, respectively. For both conditions, the judgments of the Everything-Right group (far left panels) are very close to the correct answers, with large and significant interactions of temperature × quantity, $F(16,688) = 49.71$, $F(16,672) = 100.31$ for intuitive and computational, respectively. Subjects whose patterns were midway between the Everything-Right and Everything-but-Crossover prototypes also showed significant interactions between quantity and temperature, $F(16,144) = 5.76$, $F(16,160) = 4.28$, although the crossover effect of quantity was not clear. In the intu-

the components we chose is that it would be nonsense to hypothesize a prototype such as Everything-but-Above-and-Below or Everything-but-Monotonicity because the prototypes would be outside the boundaries of the space defined by the possible component scores.

TABLE 1

NUMBERS OF SUBJECTS CLOSEST TO EACH PROTOTYPE

Prototype	Prediction Task		Reverse Task	
	Intuitive	Computational	Intuitive	Computational
(1) Everything-Right	44	43	29	34
Between (1) and (2)	10	11	6	6
(2) Everything-but-Crossover	12	5	17	8
(3) Everything-but-Range	17	2	6	2
Between (3) and (4)	5	3	4	1
(4) Everything-but-Range-and-Crossover	5	7	24	14
(5) Adding	2	20	. . .	. . .
(6) Large-Minus-Small	. . .	. . .	2	15
(7) Final-Minus-Standard	. . .	. . .	6	13
(8) Everything-Wrong	2	12	3	10

NOTE.—Subjects who were between (3) and (4) were combined with category (4) for the chi-square test. In the reverse task, one subject in the intuitive condition was closest to the Standard-Minus-Final prototype. For the chi-square test, this subject was combined with the Large-Minus-Small category.

itive condition (upper row of panels), the Everything-but-Crossover group does not show an interaction between quantity and added temperature, $F(16,176) = .76$, N.S., and the curves for all quantities are almost exactly on top of one another. In contrast, the mean judgments of subjects in the intuitive Everything-but-Range group (upper far right panel of Fig. 3) show a significant added temperature × quantity interaction, $F(16,256) = 13.75$. These subjects appear to have a solid understanding of how quantity works in the task. However, many of the judgments are out of range, that is, they do not fall between the added temperature and the 40° standard.

For the computational condition, the mean judgments in the Adding group show the expected pattern: all the means are above the 40° standard, there is no effect of or interaction with quantity, but there is a clear effect of added temperature. The mean judgments of those in the Everything-Wrong group of the computational condition (lower far right panel of Fig. 3) do not even show a significant effect of added temperature. This extremely poor understanding occurred very infrequently in the intuitive condition.

The age distribution of those individuals who are closest to each fuzzy level prototype is shown in Figure 4 with some of the

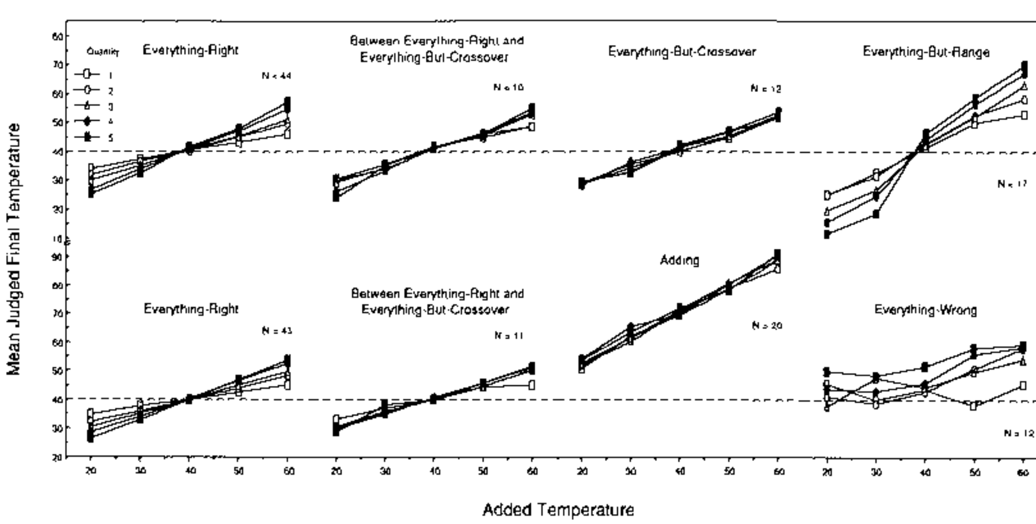

FIG. 3.—Mean judged final temperatures for the four most populous fuzzy developmental levels in the intuitive (upper panels) and computational (lower panels) conditions of the prediction task.

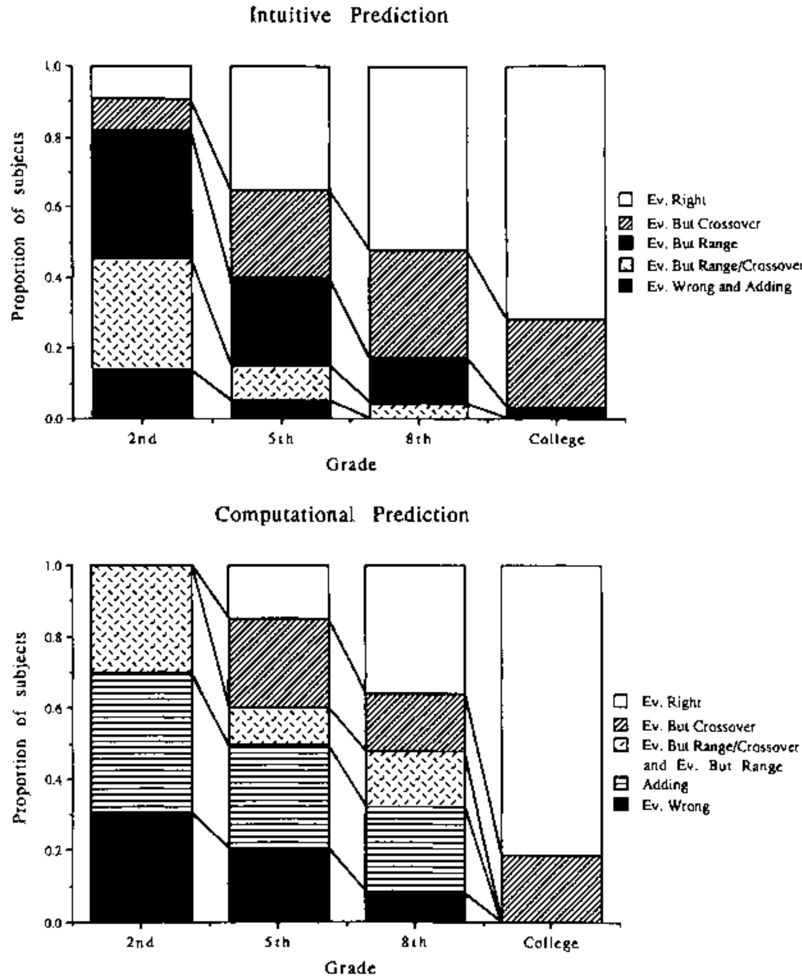

FIG. 4.—Percentage of subjects in each age group located closest to each prototype in the prediction task. Lines connecting the same prototype groups in different ages allow inspection of the developmental changes.

groups combined. In the intuitive condition (top panel), a large proportion of the college students had their highest degree of membership in the Everything-Right fuzzy level, and the proportion of subjects in Everything-Right increased with increasing age. The proportion in the Everything-but-Crossover group also increased through eighth grade. In contrast, the proportions of subjects in Everything-but-Range and Everything-but-Range-and-Crossover declined with increasing age. A chi-square test of independence on the intuitive condition was significant, $\chi^2(12) = 44.81$, $N = 97$, $p < .01$, showing that fuzzy level membership is related to age.[6]

In the computational condition (bottom panel of Fig. 4), the age trends are similar to those seen in the intuitive condition. The

[6] The chi-square test of the data in Figure 4 had a large number of cells with expected values less than 5. We also calculated the chi-square by combining the second and fifth graders and the Everything-Wrong, Adding, and Everything-but-Range-and-Crossover groups. The result was significant, $\chi^2(4) = 33.49$, $N = 97$, $p < .01$. DeLucci (1983) reported that the recommendations with respect to expected values are quite conservative with respect to Type I error. Chi-square tests for age differences in the other condition-task combinations also remained significant when categories were combined, $\chi^2(4) = 50.17, 43.64, 72.07$ for computational prediction, intuitive reverse, and computational reverse, respectively.

proportions in the Everything-Right and Everything-but-Crossover groups increase with age, while the proportions in the Adding and other fuzzy developmental levels decline with age. The chi-square test of independence on the computational condition was also significant, $\chi^2(12) = 66.40$, $N = 103$, $p < .01$. Thus, the results in Table 1 and Figure 4 show that the fuzzy level memberships for the prediction task are related to both age and condition.

Fuzzy Set Analysis of the Reverse Task

The reverse task was analyzed analogously to the prediction task. We defined eight prototype patterns and computed the Euclidean distance of each subject's component pattern to each prototype. Five of these patterns were the same as those used in the prediction task: Everything-Right, Everything-but-Crossover, Everything-but-Range, Everything-but-Crossover-and-Range, and Everything-Wrong. We replaced the Adding prototype of the prediction task with three possible subtraction patterns: Final-Minus-Standard, Larger-Minus-Smaller, Standard-Minus-Final.[7] The subtraction prototypes were used because the reverse task requires working back from the final temperature to the added temperature.

As in the prediction task, we expected that there would be a difference between the intuitive and computational conditions in the distribution of subjects across fuzzy developmental levels. The right-hand columns of Table 1 present the numbers of subjects in each condition of the reverse task closest to each prototype. As in the prediction task, some subjects were equidistant from two prototypes. A chi-square test of independence showed a significant difference in the distribution of subjects across levels for the two conditions, $\chi^2(7) = 25.70$, $N = 200$, $p < .01$. As expected, the two subtraction and the Everything-Wrong patterns had more members from the computational condition than from the intuitive condition. (Only one subject was closest to the Standard-Minus-Final prototype, so it was eliminated from further analyses.) Similarly, the Everything-but-Range and Everything-but-Range-and-Crossover patterns had more subjects from the intuitive condition closest to them. Finally, the single developmental sequence hypothesis was also contradicted in the intuitive condition of the reverse task.

Figure 5 presents the mean judgments for the reverse task for the four fuzzy level groups with the most subjects. A 4 (group) $\times$ 5 (final temperature) $\times$ 5 (quantity) analysis of variance of each condition showed significant group $\times$ final temperature $\times$ quantity interactions, $F(48,1152) = 9.61$, $F(48,1152) = 10.84$ for intuitive and computational, respectively. As in the prediction task, the mean judgments of the Everything-Right group for both the intuitive and computational conditions approximate the correct pattern with large significant interactions of temperature $\times$ quantity, $F(16,448) = 30.84$, $F(16,528) = 78.19$ for intuitive and computational, respectively. The mean judgments of those subjects in the intuitive condition who are closest to the Everything-but-Crossover prototype show no interaction between final temperature and quantity, $F(16,256) = 1.33$. In contrast, the intuitive Everything-but-Range group showed a significant interaction of final temperature $\times$ quantity, $F(16,80) = 5.62$. The mean judgments of the Everything-but-Range-and-Crossover group did not show an interaction between quantity and temperature, $F(16,368) = 2.22$, N.S., and many of their mean judgments violate the range principle. The judgments in the computational condition of the Final-Minus-Standard and Large-Minus-Small groups show patterns that differ dramatically from the others. All the means are below 40°, as expected if the subjects are literally subtracting. The Final-Minus-Standard group showed a significant main effect of final temperature, $F(4,48) = 25.39$, while the Large-Minus-Small group did not, $F(4,56) = .82$, N.S.

Figure 6 presents the percentage of subjects in each age group in each reverse task fuzzy developmental level. In both conditions of the reverse task, fuzzy level membership was significantly related to age, $\chi^2(12) = 54.62$, $N = 97$, $p < .01$, $\chi^2(12) =$

[7] The subtraction component patterns, like the adding pattern, were calculated from the response pattern that would be generated from that strategy. The Final-Minus-Standard strategy consisted of subtracting the standard from the given final temperature. The percentage component scores were 100, 100, 50, 50, 50, and 100 for main, monotonicity, above-and-below, range, crossover, and linearity. The Larger-Minus-Smaller strategy consisted of subtracting the smaller temperature from the larger one, with percentage component scores of 0, 50, 50, 50, 50, and 88.8. The Standard-Minus-Final strategy consisted of subtracting the given final temperature from the standard, and produced percentage component scores of 0, 0, 50, 50, 50, and 100.

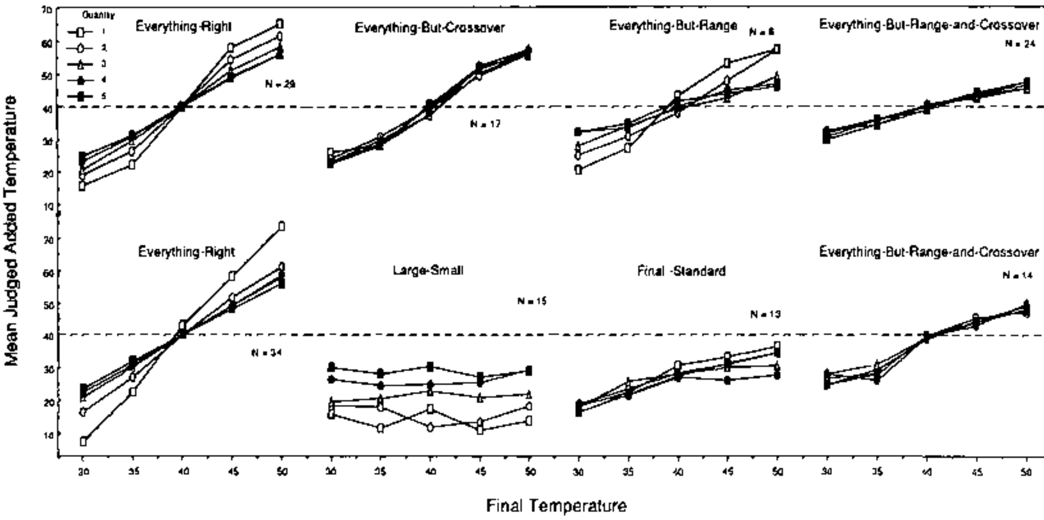

FIG. 5.—Mean judged added temperatures for the four most populous fuzzy developmental levels in the intuitive (upper panels) and computational (lower panels) conditions of the reverse task.

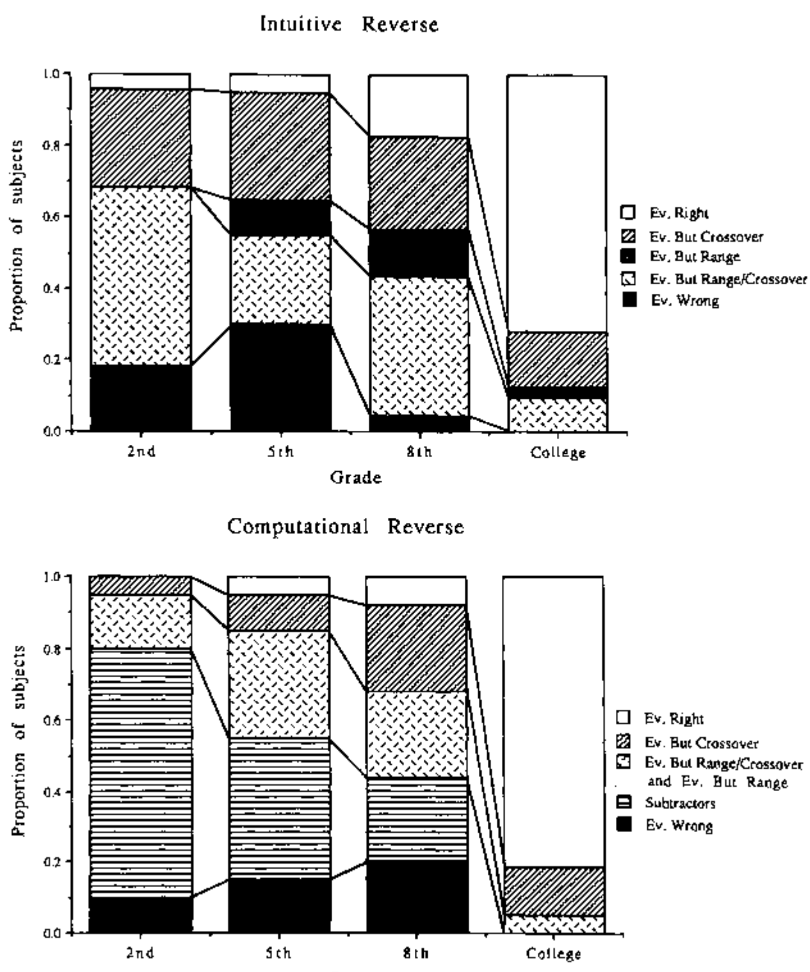

FIG. 6.—Percentage of subjects in each age group located closest to each prototype in the reverse task.

84.63, $N = 103$, $p < .01$, for intuitive and computational conditions, respectively. The majority of college students were closest to the Everything-Right prototype, and the percentage in Everything-Right increased with age. In the intuitive condition, the proportion of subjects in second through eighth grade who were in either the Everything-but-Crossover, Everything-but-Range, or Everything-but-Range-and-Crossover groups was approximately constant. In the computational condition, the proportion closest to Everything-Right increased with age, while the proportion in the Subtracting groups declined. Eighty percent of the second graders were closest to the Subtracting or Everything-Wrong prototypes. Clearly, the computational reverse task posed problems for many second graders.

Relation between Prediction and Reverse Tasks

Table 2 presents the relation between fuzzy developmental level in the prediction and reverse tasks for both the intuitive and computational conditions. First, Cohen's kappa (Siegel & Castellan, 1988) was used as a measure of consistency across tasks: $\kappa = .1461$, $z = 2.4839$, $p < .05$, and $\kappa = .4111$, $z = 7.3908$, $p < .01$ for the intuitive and computational conditions, respectively.

TABLE 2

NUMBERS OF SUBJECTS CLOSEST TO EACH PROTOTYPE

Reverse Task Prototype	Intuitive Condition, Prediction Task Prototype					
	1	2	3	4	5	8
1	19	8	2	0	0	0
2	9	8	4	1	1	0
3	4	0	2	0	0	0
4	11	5	6	5	1	0
6 & 7	1	0	3	3	0	1
8	0	1	0	1	0	1
	Computational Condition, Prediction Task Prototype					
	1	2	3	4	5	8
1	29	5	0	0	0	0
2	7	6	0	1	0	0
3	0	1	1	0	0	0
4	6	2	0	2	2	3
6 & 7	1	2	1	4	15	5
8	0	0	0	3	3	4

NOTE.—Prototypes are numbered as in Table 1. Those subjects who were between two prototypes were grouped with the higher-numbered category.

There is significant but weak consistency in the groupings across the prediction and reverse tasks. This low consistency would be expected if development in the prediction and reverse tasks is asynchronous.

The direction of the difference between the prediction and reverse tasks was tested using McNemar's change test (McNemar, 1949; Siegel & Castellan, 1988). We expected that performance would be more advanced in the prediction task than in the reverse task. For the intuitive condition, there were 23 subjects above the diagonal and 38 below the diagonal, a difference in the correct direction, but nonsignificant, $\chi^2(1) = 3.69$, $N = 61$, $p > .05$. For the computational condition, there were 16 above the diagonal and 30 below the diagonal, $\chi^2(1) = 4.26$, $N = 46$, $p < .05$. Thus, in the computational condition, subjects who were not in the same groups on the two tasks were more likely to be in a more advanced group on the prediction task than on the reverse task.

Individual Differences and Developmental Paths

There are two types of individual differences of interest to developmental researchers: (*a*) differences among individuals in the same age group, and (*b*) differences among individuals at the same developmental level. Within-age group differences are presented in Figures 4 and 6, which show that those in the same age group have different memberships in the fuzzy developmental levels.

The fuzzy set approach also provides a measure of degree of membership of each individual in each fuzzy developmental level. The percentage distance to each prototype was calculated for each subject (distance to the prototype divided by the sum of the subject's distances to all prototypes). The means and standard deviations of the percentage distances to the subjects' nearest prototypes are presented in Table 3. These values provide measures of the "fuzziness" of the groups around their prototypes. In all four conditions, the Everything-Right group is closest to its prototype. The largest mean distances occur for groups that are expected to be small or empty, suggesting that those subjects closest to those prototypes are there either by chance or through inconsistent responses. For example, the intuitive Adding and the computational Everything-but-Range prediction groups both have large means and are groups that were expected to be small or empty.

TABLE 3

MEAN PERCENTAGE DISTANCES TO NEAREST PROTOTYPE

	Intuitive			Computational		
	Mean	SD	N	Mean	SD	N
Prediction prototype:						
(1) Everything-Right	2.49	3.19	44	1.70	2.39	43
(2) Everything-but-Crossover	5.38	4.03	12	4.59	4.76	5
(3) Everything-but-Range	6.50	4.06	17	13.28	1.29	2
(4) Everything-but-Range-and-Crossover	8.12	3.55	5	6.41	4.70	7
(5) Adding	13.95	.02	2	3.94	5.14	20
(6) Everything-Wrong	9.63	1.01	2	9.97	2.97	12
Reverse prototype:						
(1) Everything-Right	1.22	1.31	29	.94	1.24	34
(2) Everything-but-Crossover	3.14	2.13	17	3.44	2.91	8
(3) Everything-but-Range	3.75	1.37	6	3.52	.35	2
(4) Everything-but-Range-and-Crossover	4.63	1.73	24	5.40	1.75	14
(5) Large-Minus-Small	7.57	2.46	6	2.88	3.04	13
(6) Final-Minus-Standard	3.90	...	1	5.13	3.03	15
(7) Everything-Wrong	6.21	.61	3	5.37	1.43	10

NOTE.—Subjects equidistant between two prototypes were omitted.

Second, in the fuzzy set approach, the rank orders of the distances to the prototypes can be used to map the developmental paths for a task. This way of displaying individual differences is shown in Figure 7. Four of the prototypes form the corners of a rectangle in a plane with the range and crossover components as its two axes. (Because the variables were divided by their SDs in calculating the distances, the axes in Fig. 7 will not form exact squares, but rectangles.) Other prototypes are omitted because they do not fall on this plane. The Everything-Wrong, Adding, and all Subtracting prototypes have component scores on variables other than range and crossover, which fall below the plane in Figure 7. The rectangle in Figure 7 can be divided into eight slices, which represent the eight possible distance rank orders in a plane (Coombs, 1964). Each slice of the rectangle shows the number of subjects who had a particular distance rank order. Only those subjects who were closest to one of these four particular prototypes are included.

Examination of Figure 7 shows that there are some subjects closest to the Everything-Right prototype who are next closest to the Everything-but-Crossover prototype, and others who are next closest to Everything-but-Range. Individuals in these two regions differ in their relative mastery of the range and crossover components even though they are all closest to the Everything-Right prototype. In the intuitive prediction task, there were only five subjects closest to Everything-but-Range-and-Crossover. However, two of them are next closest to Everything-but-Crossover, and three are next closest to Everything-but-Range. These could represent the beginnings of the two alternative developmental paths. Those subjects in the 3412 ($N = 8$) and 2413 ($N = 6$) slices are the continuations of the two paths. Thus, an individual's distance rank order provides a simplified profile of position with respect to the hypothesized fuzzy developmental level prototypes, and characterizes the person's current state of development.

Discussion

Proportional reasoning can be conceptualized in terms of a set of components that develop gradually through several organizational patterns or structures. There are two major contingencies on the structure of components: (*a*) whether the task is presented numerically with computation encouraged, or such that only estimation and intuition are required, and (*b*) whether the task is presented such that it requires prediction versus inference of a prior state (reverse task). Even a fair number of fifth graders showed mature proportional reasoning in the intuitive version of the temperature prediction task, in contrast to the reverse task or numerical prediction task. These findings provide a new perspective on proportional reasoning.

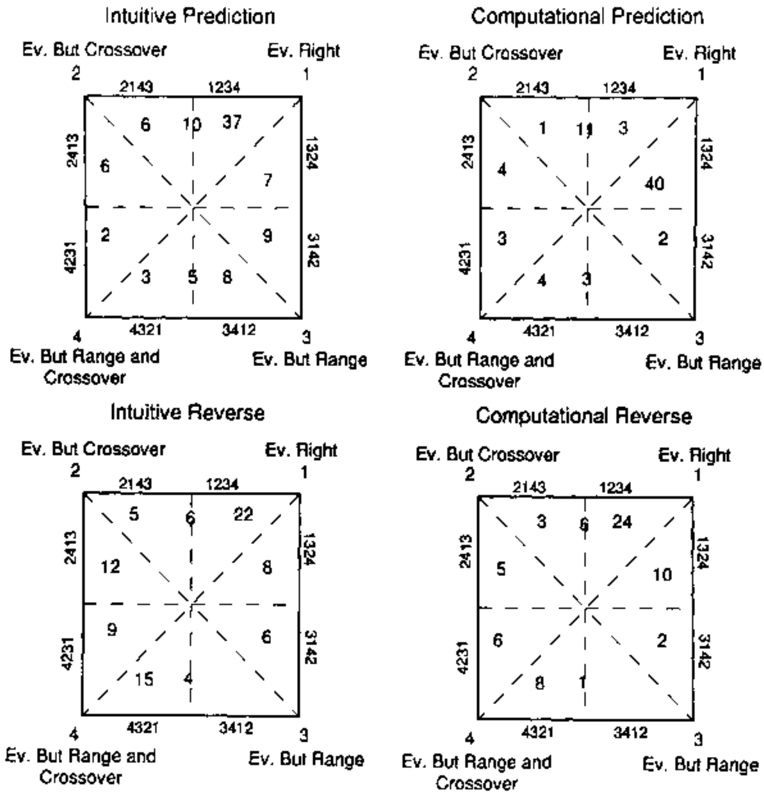

FIG. 7.—Frequencies of individuals in regions of the plane described by four of the developmental level prototypes: (1) Everything-Right, (2) Everything-but-Crossover, (3) Everything-but-Range, and (4) Everything-but-Range-and-Crossover. The sequences of numbers around the edge correspond to the rank order of distances to the prototypes for each region.

Intuitive versus Numerical Proportional Reasoning

Inhelder and Piaget (1958, p. 219) viewed proportional reasoning, which involves explicit calculations, as somehow following naturally from qualitative proportional reasoning. The present research shows that it is inadequate to describe numerical proportional reasoning as automatically arising from mature intuitive proportional reasoning. The relation between intuitive understanding and the process of arriving at a computational scheme deserves further research. First, the ability of subjects to use their intuitive understanding of the task to evaluate different computational formulas needs to be studied. For example, subjects with knowledge of all components except range would not be expected to be satisfied with computing a simple average of temperatures, because doing so ignores quantity. Second, the relation between intuitive understanding and analogical reasoning should be studied. Even if a person has the correct weighted average formula or the intuitive weighted average concept in another task, it must still be mapped correctly onto the temperature mixture task (Gentner, 1983). A person's intuitive understanding of the temperature task should be important in the process of mapping relations in one domain onto another domain because mapping requires some understanding of how variables function in a problem. In sum, unless a person attempts to bring intuitive understanding to bear on a computational task, a computational scheme would be expected to be based primarily on the memory availability of mathematical operations. This is essentially what is seen when subjects add or subtract the temperatures. Further research is needed to explore the conditions under which subjects do and do not use their intuitive understanding to regulate their computational attempts.

Proportional Reasoning on Prediction versus Reverse Task

A second major contingency is whether the task is presented in a prediction format or inference (reverse task) format. The fuzzy set analysis and mean component scores show that the components of understanding are less firmly grasped in the reverse task. Other research has shown that problems that involve the same mathematical operations can differ greatly in difficulty. Greer (1987) reviewed the literature and concluded that addition and subtraction problems in which the final state is unknown are easier than those in which one of the other variables is unknown. For multiplication and division problems, however, Greer concluded that the situation is much more complex, with the type of number (integer, decimal greater than 1, or decimal less than 1) having a very large influence on difficulty. In the present study, the children performed worse on the reverse task compared to the prediction task in *both* the computational and intuitive conditions. The prediction and reverse tasks may involve different logic or sets of operations. This interpretation is supported by research on social judgments using prediction and reverse tasks (Kun, 1977; Surber, 1980, 1984) and other research on physical tasks (Surber & Gzesh, 1984; Wilkening, 1981).

One possible explanation of the difficulty of the reverse task is that it requires reasoning about a hypothetical prior state, the type of reasoning that might be said to require formal operations since it involves supposing that the containers had not actually been combined (see Kun, 1977). It is also possible that the working memory load involved in formulating answers for the reverse task is higher than that of the prediction task. If a strategy for a task is so complex that it overloads working memory, the subject will either show errors in executing the strategy (and thus perform more poorly) or revert to a simpler (and less adequate) strategy with lower requirements on working memory (Shatz, 1978). A drawback of the working memory explanation is that it assumes that over a wide age range people are capable of using a variety of strategies, and that what develops is memory capacity, and the automaticity of the strategies which allows them to be applied with limited memory capacity. Thus, the process of the development of new knowledge structures or strategies is de-emphasized. Based on the present results, it seems that in order to have a full account of the development of proportional reasoning, some process of computational strategy invention (Siegler & Jenkins, 1989) or intuitive knowledge restructuring needs to be included.

Modeling Developmental Change: Components and "Fuzzy Sets"

In age group analyses, the components showed different developmental functions and different effects of the experimental manipulations. However, age group analyses ignore individual differences and can actually distort the data patterns of the individuals (see Surber, 1980). The fuzzy set approach introduced here provides a way of analyzing cross-sectional developmental change which preserves individual differences. Individuals are characterized according to the similarity of their component profiles to prototypes. The prototypes can be chosen to represent qualitatively different developmental levels of performance. The developmental trends in the hypothesized fuzzy levels can be seen by examining the age distributions of those individuals with similar component profiles. The applicability of the present approach to other domains is shown in a study of perspective taking that used a similar approach. Dixon and Moore (1990) defined four components in a perspective-taking task, and grouped subjects according to their component profiles. The groups showed qualitative developmental differences that were strikingly different from the age group means, as in the present study.

The fuzzy set approach also provides a method for empirically testing the hypothesis that there is a single path of development or a universal invariant sequence. The distances of a subject to each prototype indicate where that subject is with respect to the components and the prototypes. The data of the intuitive condition of the present study were inconsistent with the universal invariant sequence hypothesis for acquisition of the crossover and range components. Perhaps cognitive developmental research has concentrated too much on the universal invariant sequence idea while neglecting the description of multiple developmental paths. The fuzzy set approach provides a way of measuring and describing the developmental paths that exist in a domain.

Although individuals can be classified discretely as members of a particular fuzzy developmental level, the approach represents individuals as having different degrees of membership in developmental levels. This provides a potential rapprochement between continuity theories of development and theories that emphasize qualitative changes. The question of whether development is a continuous process of quantitative change as opposed to a process of shifts between qualitatively different levels then becomes an issue of grain of analysis, or scale in systems theory (Allen & Starr, 1982). If developmental levels are categories with fuzzy boundaries, then the rate of change between levels can be measured. Once such measurements are possible, researchers then can specify how rapidly a change from one level to the next must occur in order for it to be considered a saltatory, qualitative shift. Thus, traditional classification methods in which individuals are placed in discrete developmental levels are incomplete, but not necessarily incorrect. The fuzzy set approach allows simultaneous description of individual differences, developmental levels, and the pathways of development.

References

Allen, T. F. H., & Starr, T. B. (1982). *Hierarchy*. Chicago: University of Chicago Press.

Anderson, N. H., & Butzin, C. A. (1978). Integration theory applied to children's judgments of equity. *Developmental Psychology*, **14**, 593–606.

Brainerd, C. J. (1973). Order of acquisition of transitivity, conservation, and class inclusion of length and weight. *Developmental Psychology*, **8**, 105–116.

Brainerd, C. J. (1977). Response criteria in concept development research. *Child Development*, **48**, 360–366.

Coombs, C. H. (1964). *A theory of data*. New York: Wiley.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, **12**, 671–684.

DeLucci, K. L. (1983). The use and misuse of chi-square: Lewis and Burke revisited. *Psychological Bulletin*, **94**, 166–176.

Dixon, J. A., & Moore, C. F. (1990). The development of perspective taking: Understanding differences in information and weighting. *Child Development*, **61**, 1502–1513.

Ferretti, R. P., Butterfield, E. C., Cahn, A., & Kerkman, D. (1985). The classification of children's knowledge: Development on the balance scale and inclined plane tasks. *Journal of Experimental Child Psychology*, **39**, 131–160.

Flavell, J. H. (1971). Stage-related properties of cognitive development. *Cognitive Psychology*, **2**, 421–453.

Frankel, P., & Strauss, S. (1985). Children's understanding of temperature as assessed by three developmental models. *SRCD Biennial Meeting Abstracts*, **5**, 112.

Gentner, D. (1983). Structure-mapping: A theoretical framework. *Cognitive Science*, **7**, 155–170.

Greer, B. (1987). Understanding of arithmetical operations as models of situations. In J. A. Sloboda & D. Rogers (Eds.), *Cognitive processes in mathematics* (pp. 60–80). Oxford: Clarendon Press.

Hammond, K. R., Hamm, R. M., Grassia, J., & Pearson, T. (1987). Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment. *IEEE Transactions*, SMC-17, 753–770.

Hook, J. G., & Cook, T. D. (1979). Equity theory and the cognitive ability of children. *Psychological Bulletin*, **86**, 429–445.

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

Jackson, E., Campos, J. J., & Fischer, K. W. (1978). The question of decalage between object permanence and person permanence. *Developmental Psychology*, **14**, 1–10.

Kerkman, D. D., & Wright, J. C. (1988). An exegesis of two theories of compensation development: Sequential decision theory and information integration theory. *Developmental Review*, **8**, 323–360.

Kermoian, R., & Campos, J. J. (1988). Locomotor experience: A facilitator of spatial cognitive development. *Child Development*, **59**, 908–917.

Kessen, W. (1960). Research design in the study of developmental problems. In P. H. Mussen (Ed.), *Handbook of research methods in child development* (pp. 36–70). New York: Wiley.

Kun, A. (1977). Development of the magnitude-covariation and compensation schemata in ability and effort attributions of performance. *Child Development*, **48**, 862–873.

Kun, A., Parsons, J., & Ruble, D. (1974). Development of integration processes using ability and effort information to predict outcome. *Developmental Psychology*, **10**, 721–732.

Langer, J. (1970). Werner's comparative organismic theory. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology*, 3d ed. (Vol. **1**, pp. 733–771). New York: Wiley.

McNemar, Q. (1949). *Psychological statistics*. New York: Wiley.

Nicholls, J. G., & Miller, A. T. (1984). Development and its discontents: The differentiation of the concept of ability. In J. G. Nicholls (Ed.), *The development of achievement motivation* (pp. 185–218). Greenwich, CT: JAI.

Noelting, G. (1980a). The development of proportional reasoning and the ratio concept: I. Differentiation of stages. *Educational Studies in Mathematics*, **11**, 217–253.

Noelting, G. (1980b). The development of proportional reasoning and the ratio concept: II. Problem-structure at successive stages, problem-solving strategies and the mechanism of adaptive restructuring. *Educational Studies in Mathematics*, **11**, 331–363.

Norman, K. L., & Schemmer, F. M. (1977). *Functional processing in information integration tasks* (Center for Language and Cognition, Tech. Rep. No. 14). University of Maryland.

Oden, G. C. (1977). Integration of fuzzy logical information. *Journal of Experimental Psychology: Human Perception and Performance*, **3**, 565–575.

Reed, S. K., & Evans, A. C. (1987). Learning functional relations: A theoretical and instructional analysis. *Journal of Experimental Psychology: General*, **116**, 106–118.

Selman, R. L., & Byrne, D. F. (1974). A structural developmental analysis of levels of role taking in middle childhood. *Child Development*, **45**, 803–806.

Shatz, M. (1978). The relationship between cognitive processes and the development of communication skills. In C. B. Keasey (Ed.), *Nebraska symposium on motivation 1977* (Vol. **25**, pp. 1–42). Lincoln: University of Nebraska Press.

Siegel, S., & Castellan, N. J., Jr. (1988). *Nonparametric statistics for the behavioral sciences* (2d ed.), New York: McGraw-Hill.

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology*, **8**, 481–520.

Siegler, R. S. (1981). Developmental sequences within and between concepts. *Monographs of the Society for Research in Child Development*, **46**(2, Serial No. 189).

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, **116**, 250–264.

Siegler, R. S. (1988). Strategy choice procedures and the development of multiplication skill. *Journal of Experimental Psychology: General*, **117**, 258–275.

Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, NJ: Erlbaum.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.

Strauss, S., & Stavy, R. (1982). U-shaped behavioral growth: Implications for theories of development. In W. W. Hartup (Ed.), *Review of child development research* (Vol. **6**, pp. 547–599). Chicago: University of Chicago Press.

Surber, C. F. (1980). The development of reversible operations in judgments of ability, effort, and performance. *Child Development*, **51**, 1018–1029.

Surber, C. F. (1984). Inferences of ability and effort: Evidence for two different processes. *Journal of Personality and Social Psychology*, **46**, 249–268.

Surber, C. F., & Gzesh, S. M. (1984). Reversible operations in the balance scale task. *Journal of Experimental Child Psychology*, **38**, 254–274.

Surber, C. F., & Haines, B. A. (1987). The growth of proportional reasoning: Methodological issues. In R. Vasta & G. Whitehurst (Eds.), *Annals of Child Development* (Vol. **4**, pp. 35–87). Greenwich, CT: JAI.

Uzgiris, I., & Hunt, J. (1975). *Assessment in infancy: Ordinal scales of psychological development*. Urbana: University of Illinois Press.

Wilkening, F. (1981). Integrating velocity, time, and distance information: A developmental study. *Cognitive Psychology*, **13**, 231–247.

Wilkening, F. (1988). A misrepresentation of knowledge representation. *Developmental Review*, **8**, 361–367.

Wilkinson, A. C., & Haines, B. A. (1987). Learning a cognitive skill and its components. In J. Bisanz, C. J. Brainerd, & R. V. Kail (Eds.), *Formal methods in developmental psychology* (pp. 86–114). New York: Springer-Verlag.

Wohlwill, J. F. (1973). *The study of behavioral development*. New York: Academic Press.

Zadeh, L. (1965). Fuzzy sets. *Information and Control*, **8**, 338–353.

LINKED CITATIONS

Components of Understanding in Proportional Reasoning: A Fuzzy Set Representation of Developmental Progressions

Colleen F. Moore; James A. Dixon; Beth A. Haines

Child Development, Vol. 62, No. 3. (Jun., 1991), pp. 441-459.

Stable URL:

http://links.jstor.org/sici?sici=0009-3920%28199106%2962%3A3%3C441%3ACOUIPR%3E2.0.CO%3B2-D

References

Response Criteria in Concept Development Research

Charles J. Brainerd

Child Development, Vol. 48, No. 2. (Jun., 1977), pp. 360-366.

Stable URL:

http://links.jstor.org/sici?sici=0009-3920%28197706%2948%3A2%3C360%3ARCICDR%3E2.0.CO%3B2-K

The Development of Perspective Taking: Understanding Differences in Information and Weighting

James A. Dixon; Colleen F. Moore

Child Development, Vol. 61, No. 5. (Oct., 1990), pp. 1502-1513.

Stable URL:

http://links.jstor.org/sici?sici=0009-3920%28199010%2961%3A5%3C1502%3ATDOPTU%3E2.0.CO%3B2-V

Locomotor Experience: A Facilitator of Spatial Cognitive Development

Rosanne Kermoian; Joseph J. Campos

Child Development, Vol. 59, No. 4. (Aug., 1988), pp. 908-917.

Stable URL:

http://links.jstor.org/sici?sici=0009-3920%28198808%2959%3A4%3C908%3ALEAFOS%3E2.0.CO%3B2-U

LINKED CITATIONS

Development of the Magnitude-Covariation and Compensation Schemata in Ability and Effort Attributions of Performance
Anna Kun
Child Development, Vol. 48, No. 3. (Sep., 1977), pp. 862-873.
Stable URL:
http://links.jstor.org/sici?sici=0009-3920%28197709%2948%3A3%3C862%3ADOTMAC%3E2.0.CO%3B2-P

A Structural-Developmental Analysis of Levels of Role Taking in Middle Childhood
Robert L. Selman; Diane F. Byrne
Child Development, Vol. 45, No. 3. (Sep., 1974), pp. 803-806.
Stable URL:
http://links.jstor.org/sici?sici=0009-3920%28197409%2945%3A3%3C803%3AASAOLO%3E2.0.CO%3B2-B

Developmental Sequences within and between Concepts
Robert S. Siegler; Sidney Strauss; Iris Levin
Monographs of the Society for Research in Child Development, Vol. 46, No. 2, Developmental Sequences within and between Concepts. (1981), pp. 1-84.
Stable URL:
http://links.jstor.org/sici?sici=0037-976X%281981%2946%3A2%3C1%3ADSWABC%3E2.0.CO%3B2-O

The Development of Reversible Operations in Judgments of Ability, Effort, and Performance
Colleen F. Surber
Child Development, Vol. 51, No. 4. (Dec., 1980), pp. 1018-1029.
Stable URL:
http://links.jstor.org/sici?sici=0009-3920%28198012%2951%3A4%3C1018%3ATDOROI%3E2.0.CO%3B2-U