DOCUMENT RESUME

ED 238 708 SE 043 684

AUTHOR	Shaklee, Harriet; Paszek, Donald
TITLE	Eliciting Systematic Rule Use in Covariation Judgment [the Early Years].
INSTITUTION	Iowa Univ., Iowa City. Dept. of Psychology.
SPONS AGENCY	National Inst. of Education (ED), Washington, DC.
PUB DATE	[83]
CONTRACT	NIE-G-80-0091
NOTE	35p.; Appendix C of SE 043 682.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Cognitive Processes; Educational Research; Elementary Education; *Evaluative Thinking; Individual Differences; *Learning; Mathematical Concepts; *Mathematics Education; Prediction; *Probability; Problem Solving; *Psychological Studies
IDENTIFIERS	*Mathematics Education Research; Rule Learning

ABSTRACT

Related research suggests that children may show some simple understanding of event covariations by the early elementary school years. The present experiments use a rule analysis methodology to investigate covariation judgments of children in this age range. In Experiment 1, children in second, third and fourth grade judged covariations on 12 different covariation problems. Children's performance patterns on the problem set showed an increase in the use of systematic judgment strategies in this age range. Systematic rule users most commonly compared contingency table cells "a" and "b" in judging the event covariations. In Experiment 2, a training paradigm was employed to investigate possible origins of systematic rule use. First and second grade unsystematic, strategy 0 and cell-a children were either directed to attend to cells "a" and "b" (Attention only), were additionally offered explicit instructions to note which of the two cells had more events (Attention-plus-more) or were given no training (control). Posttest performance showed that the Attention-plus-more condition was the only treatment to reliably elicit a-versus-b rule use. It is concluded that simple covariation judgment rules can be used by children in the early elementary school years. (Author)

Eliciting Systematic Rule Use in Covariation Judgment [in Early Years]

Harriet Shaklee and Donald Paszek

University of Iowa

Running head: Covariation Judgment

Partial support for this research was provided by NIE grant NIE-G-80-0091. Many thanks to Ernest Jones, Renee Smith, Rick Taffe and Janet Lyness for their help with data collection. Reprint requests should be sent to Harriet Shaklee, Department of Psychology, University of Iowa, Iowa City, Iowa 52242.

Abstract

Related research suggests that children may show some simple understanding of event covariations by the early elementary school years. The present experiments use a rule analysis methodology to investigate covariation judgments of children in this age range. In Experiment 1, children in second, third and fourth grade judged covariations on 12 different covariation problems. Children's performance patterns on the problem set showed an increase in the use of systematic judgment strategies in this age range. Systematic rule users most commonly compared contingency table cells a and b in judging the event covariations. In Experiment 2, a training paradigm was employed to investigate possible origins of systematic rule use. First and second grade unsystematic, strategy 0 and cell-a children were either directed to attend to cells a and b (Attention only), were additionally offered explicit instructions to note which of the two cells had more events (Attention-plus-more) or were given no training (control). Posttest performance showed that the Attention-plus-more condition was the only treatment to reliably elicit a-versus-b rule use. It is concluded that simple covariation judgment rules can be used by children in the early elementary school years.

Children's Judgments about Covariation between Events:
A Series of Training Studies. Appendix C.

Covariation Judgment: Systematic Rule Use in the Early Years

Interest in children's causal reasoning has burgeoned in recent years (e.g., Siegler, 1976; Bullock, Gelman & Baillargeon, 1982). A number of theorists have suggested that identification of cause-effect relationships is grounded in covariation judgment (e.g., Inhelder & Piaget, 1958; Kelley, 1972). That is, people search for causes of events by finding event covariates. In fact, a few investigations indicate that children understand this link from an early age. For example, DiVitto and McArthur (1978) found that children as young as first grade use summarized covariation information in explaining people's behavior. Siegler and Liebert (1975), however, found that children were not influenced by event covariation until 8 or 9 years of age in their study of children's explanations of physical events. Evidence of the earliest use of event covariation in causal reasoning is provided by Shultz and Mendelson (1975), who found that 3 and 4 year old children showed a preference for covariates when choosing causes of events. Although the age trends differ in these studies, they concur in suggesting that preference for consistent covariates is an early developing pattern in children's explanations of events.

Given this evidence, understanding development in covariation judgment would be critical to understanding children's causal reasoning. However, investigations of children's abilities to make covariation judgments are rare indeed. Those few studies which do exist show a degree of consensus on how children might judge event relationships (Inhelder & Piaget, 1958; Adi, Karplus, Lawson & Pulos, 1978; Shaklee & Mims, 1981). In the basic paradigm investigators offered subjects information on the frequency of cooccurrence of alternative event states of two potentially related variables (for example, plants healthy or not healthy; plant food present or absent). Subjects were asked to identify the direction and/or strength of the relationship between the events. In each experiment, subjects' covariation judgments and/or explanations of those judgments led the investigators to identify systematic but inaccurate rules which were precursors to the use of more mathematically sophisticated rules.

Inhelder and Piaget (1958) proposed two simple rules of covariation judgment. In the first, an individual would judge a relationship according to the frequency with which target event states cooccur (e.g., healthy plants which are given plant food in the example above, cell _a_ of a traditionally labeled contingency table. See Table 1). A subject using this strategy would

Insert Table 1 here

identify a positive relationship between events if the cell _a_ frequency were the largest of the contingency table cells, and a negative relationship if it were the smallest (cell-a strategy). Inhelder and Piaget (1958) identified this strategy as common among younger adolescents. Smedslund (1963) and Nisbett and Ross (1980) thought the strategy might typify adult reasoning as well.

Also proposed by Inhelder and Piaget (1958) was a second simple approach comparing the number of times the target outcome occurs with the supposed cause (or covariate) with the number of times it occurs without that cause (for example, healthy plants with plant food vs. healthy plants without plant food). This would compare contingency table cells _a_ and _b_ (strategy a-versus-b). This strategy was identified by Inhelder and Piaget (1958) as typical of young adolescents and was found by other investigators to be common among high school subjects as well (Adi, Karplus, Lawson and Pulos, 1978).

Inhelder and Piaget (1958) proposed a third strategy as characteristic of formal operational thinking. That is, subjects would compare frequencies of events confirming (cells _a_ and _d_) and disconfirming (cells _b_ and _c_) a relationship of a particular direction. This rule would compare the sums of diagonal cells in the contingency table (sum of diagonals strategy).

Finally, Jenkins and Ward (1965) propose that covariation is most accurately assessed by comparing the conditional probabilities of an event occurring given each of the alternative states of the other variable (e.g., plant health/plant food vs. plant health/no plant food). This would compare the frequency ratio in contingency table cells $\frac{a}{a + c}$ with that in cells $\frac{b}{b + d}$ (conditional probability strategy).

This analysis of possible rules may allow diagnosis of strategies actually employed by children of various ages. That is, different rules should produce different judgments on carefully constructed covariation problems. A set of such problems is illustrated in Tables 2a and 2b. Solution accuracy is indexed

---

Insert Tables 2a and 2b here

---

by the direction of the judged relationship (i.e. A_1 more likely given B_1, B_2, or no difference). Problems are structured hierarchically such that cell-a problems are correctly solved by all strategies, a-versus-b problems are accurately solved by all strategies except cell-a. Sum-of-diagonals problems are accurately judged by sum-of-diagonals and conditional probability strategies and conditional probability problems are accurately judged by the conditional probability rule alone (see Table 3). The probability of matching these

Insert Table 3 here

judgment patterns by chance alone is .11 for cell-a, .04 for a-versus-b, .01 for sum of diagonals, and .005 for the conditional probability pattern.

Shaklee and Mims (1981) used this rule diagnostic approach to study covariation judgment strategies used by subjects from 4th grade through college age. Subjects' judgment patterns in that age span showed a strong developmental trend, with the a-versus-b strategy evidenced by substantial numbers of subjects beginning in the fourth grade (29%), and sum of diagonals the modal strategy at 7th and 10th grade (50% of subjects). Conditional probability patterns were produced by many subjects at the 10th grade (27%) but were still used by a minority of subjects even in the college years (38%). Thus, this evidence supports previous investigators' suggestions that children may use simpler, less accurate rules as precursors to mature reasoning. However, these results deviated from previous conclusions in two notable ways. First, the commonly proposed cell-a judgment pattern was rare among subjects at any of the ages tested (0-8%). In addition, the level of mature reasoning most often fell short of the optimal judgment strategy.

These results further contrast with findings in the causal reasoning research where use of covariation information was seen in causal judgment anywhere from preschool to 8-9 years of age. Shaklee and Mims (1981), on the other hand, find that nearly half of fourth graders showed no systematic bases of covariation judgment. A look at the causal reasoning research indicates that these studies offered children a relatively easy task of covariation judgment. DiVitto and McArthur (1978), for example, summarized the covariation information for the subjects, allowing children to use the information in causal judgment when they might not be able to derive that information for themselves.

In the remaining studies (Shultz & Mendelson, 1975; Siegler & Liebert, 1975), the target event and its possible causes were either perfectly contingent or completely independent. Studies of covariation judgment, on the other hand, commonly ask for judgments about less-than-perfect relationships. This analysis would indicate that young children may evidence a very simple understanding of covariation which does not hold up well when judging relationships of intermediate strength.

A final related paradigm must also be considered in understanding children's covariation judgment. That is, one commonly employed test of probability judgment is one in which a child is shown two piles of marbles composed of different proportions of marbles of two colors. The subject is asked to indicate the pile from which he or she would rather make a blind choice in order to obtain the marble of a particular color. The judgment is formally comparable to a covariation judgment, where a subject decides if a given outcome is more likely under condition A_1 or A_2. Siegler's (1981) rule analysis of children's performance in this paradigm shows systematic rule use by a narrow majority of 5 year olds with most of those children using a rule comparable to the a-versus-b rule in covariation judgment research. By 8-9 years of age a substantial majority of children were using systematic judgment rules, with a comparison of conditional probabilities the modal response pattern in Experiment 1, a-versus-b the dominantly used rule in Experiment 2. Each experiment found a comparison of conditional probabilities to be the most common rule among 12 year olds and adults.

Thus, in contrast to covariation judgment research, Siegler found that systematic rule use in a related judgment occurs at an earlier age, culminating in use of the optimal rule by early adolescence. Siegler's (1981) findings may suggest that Shaklee and Mims (1981) provide a conservative estimate of children's acquisition of systematic bases of covariation judgment. Causal reasoning research also indicates that some simple understanding of event covariation may be seen by the early elementary school years. Possible resolution of these differences may begin with a careful look at the covariation judgment paradigm. The reliable strategy use evidenced by older subjects clearly indicates that they understood the experimental stimuli and procedures. However, among the fourth grade sample, 25% of the subjects produced unclassifiable response patterns, and an additional 21% passed no strategy criteria at all. This high rate of unsystematic responses may indicate that a substantial group of these children were confused by the paradigm and thus, unable to demonstrate systematic rules which may be in their repertoires. If this were the case, a simplified approach should be developed to test these younger subjects.

We address the question of early covariation judgment in two ways. Experiment 1 employs a simplified paradigm to examine the development of covariation judgment rule use among young elementary school children. Once these normative trends are established our second study investigates sources of this shift to systematic rule use. In Experiment 2, we test information components which may be sufficient to elicit reliable rule use among young children.

Experiment 1

Simplification of our previous experimental procedure was accomplished in two major ways. First, we were concerned that younger subjects might not understand the stimuli represented in the 2 x 2 table. As a result, a new introduction expanded the discussion of the contents of the table, asking the subject to point to examples of each of the four possible combinations of event states in the table.

Secondly, we suspected that our previous question format might be overly complex for the younger children. The previous question asked (in the plant food example discussed above).

When they got special food, plants were

a) more likely to be healthy than

b) just as likely to be healthy as

c) less likely to be healthy than

when they didn't get special food.

A reformulated question offered simpler syntax:

Plants were more likely to get better if

a) they got the special food

b) they did not get the special food

c) no difference

We expected that this simplified question would be more appropriate to the language competencies of younger subjects. Experiment 1 also included two different problem sets in anticipation of our needs in the subsequent study.

Method

Subjects

Subjects in the experiment were respondents to an advertisement in a small town newspaper offeri payment to second, third and fourth grade children for participating in a psychology experiment. The resultant sample included 37 second graders, 18 third graders, and 17 fourth graders.

Problems

Subjects judged one of two sets of 12 covariation problems, each structured to produce a distinctive pattern of solution accuracy by each of the four proposed judgment rules. In one set of problems, cell frequencies totaled 36 for each problem (set 24), in the other set, cell frequencies totaled 36 for each problem (set 36). Except for these frequency differences, the two problem sets were identical in other respects. Tables 2a and 2b show the actual problem frequencies used for the problems in each of the two sets. The 12 problems in each set included three problems for each of the four strategy types. One noncontingent (middle row Tables 2a and 2b) and two contingent relationships (top and bottom rows Tables 2a and 2b, $P(A_1/B_1) - P(A_1/B_2) = (.40$ to $.50)$

were included for each problem strategy type. Table 3 shows the pattern of solution accuracy congruent with each of the proposed rules.

Each problem was set in a concrete context of two everyday events which may or may not be related. Each individual event pairing was illustrated with a small picture showing the state of the two variables (e.g., plant sick or healthy/plant food present or absent). Three problems pictured bakery products which either rose or fell in association with the presence or absence of yeast, baking powder, or a "special ingredient". In three other problems, plants were pictured as healthy or sick as a possible function of the presence or absence of plant food, bug spray, or a "special medicine". In three problems people or animals were pictured as sick or healthy as a possible function of the presence or absence of a shot, liquid medicine, or a pill. The three remaining problems pictured a possible association between space creatures appearing happy or sad in the presence or absence of one of three weather conditions (snow, fog, or sunshine).

For each problem, data instances were organized in a 2 x 2 table. In each case, the manipulated factor (or environmental event) defined the table columns (e.g., plant food, no plant food in example below), and the outcomes defined the table rows (e.g., plants healthy, not healthy in the example below). Each problem was introduced with a paragraph describing a context in which several observations were made on two potentially related variables. Subjects were asked to look at the pictured information and to identify the relative likelihood of one of the events when the second event was either present or absent. An example problem follows:

A plant grower had a bunch of sick plants. He gave some of them special plant food, but some plants didn't get special food. Some of the plants got better but some of them didn't. In the picture you will see how many times these things happened together. The picture shows that the plants were more likely to get better if:

A. they got the special food.

B. they did not get the special food.

C. no difference (they were just as likely to get better with food as without the food).

The 12 problems were grouped into problem blocks, including one problem from each strategy type. Problems within each block were arranged in a single random sequence. The three problem blocks were sequenced in a single random order. Numbers in parentheses to the left of the problems in Tables 2a and 2b indicate the position of each problem in the problem sequence.

Procedure

Each subject was tested individually. Introductory instructions introduced the subject to the concept of covariation in the context of "things that go together". Naturally occurring examples were given of positive relationships (i.e., tall people are more likely to be heavy than short people), negative relationships (i.e., it is less likely to rain when it is sunny than when it is cloudy), and unrelated events (i.e., a green truck is just as likely to run out of gas as a red truck). Subjects were told that they would be given some problems about hypothetical events that may or may not tend to go together. Two sample problems were used to clarify the information in the 2 x 2 table. The first sample problem was read to the subject. The subject was told that pictures in the cells showed the occurrence or nonoccurrence of the two events in the story. The experimenter then pointed out that each cell represented a different combination of the two possible events and stated what these were. The subject was asked to point to cells corresponding to specific combinations of events given by the experimenter. The experimenter explained that each picture in the cells represented one occurrence of a particular combination of events, so that the number of pictures in each cell represented the number of times that combination occurred. The experimenter then read the covariation question to the subject and asked him or her to answer it based on the events pictured in the table. It was emphasized that subjects should answer the questions based on what had occurred in each story problem and should avoid basing answers on knowledge of common everyday occurrences (for example, that it is more likely to snow when it is cold, regardless of cell frequencies). Each subject gave a solution to the problem and repeated the procedure on the second sample problem. Subjects were encouraged to ask any questions they might have about the task.

The subject then proceeded to the 12 problem set. Each of the problems in the set were read to the subject by the experimenter. Subjects were allowed to answer the problems at their own pace.

Results

Our main interest in this study was to establish trends in strategy use among these younger subjects. As a result, the analyses in this study use subject strategy classification as the dependent variable of interest. Subjects were classified for strategy use according to the method illustrated in Table 3. A subject was said to have "passed" a given problem type if he or she was accurate on two or more of the three problems of a given problem type.

A subject who met this criterion on all problem types would be classified as a conditional probability rule user, subjects who passed criteria on all types except the conditional probability problems were labeled sum-of-diagonals judges. A-versus-b judges should pass the cell-a and a-versus-b problems, but not the other problem types, cell-a rule users should pass criterion on cell-a problems alone. Subjects who passed no problem types were labeled Strategy 0; all other judgment patterns were categorized as unclassifiable. Table 4 shows the rule classifications of subjects in each of the three grades.

Insert Table 4 here

The modal classification at each of the grades was a-versus-b, with very few subjects showing evidence of use of more sophisticated rules and a few subjects at each grade with cell-a rule judgment patterns. Many subjects in the second and third grades made judgments that were not classifiable by any of our rules. Effects of grade level and problem set were examined by assigning subjects a score according to the number of problem type criteria passed. Thus, Strategy 0 subjects were assigned a score of 0, conditional probability subjects a score of 4. Unclassifiable subjects could not be clearly ranked in this way and were excluded from these analyses. Data from the remaining subjects were analyzed in an analysis of variance with subject's grade (2, 3, or 4) and problem set (24 or 36) as factors. These analyses showed a significant effect of grade, $\underline{F}$ $(2,51) = 3.30$, $p < .05$, with third and fourth graders similar to each other, and classified as using more advanced rules than the 2nd graders (Duncan's multiple range test, $p < .05$). Problem set effects were not significant.

Discussion

Related research in causal reasoning and probability judgment indicated that children might show some simple understanding of event covariation by early elementary school. This experiment found that a majority of children do show systematic rule use in covariation judgment by the second grade. Significant age trends also show an increase in systematic rule use with age in the second to fourth grade age span. Rule categorizations in this age range show a substantial decline in unclassifiable and Strategy 0 subjects with increasing age and an increase in a-versus-b rule use. However, use of more advanced rules was rare at all ages tested.

Comparison with Shaklee and Mims (1981) indicates that subjects did indeed show earlier competencies with our revised procedure. Nearly all fourth graders were classifiable by one of our proposed rules in the present experiment and a majority of children showed systematic rule use in the second and third grades. Overwhelmingly, these children were classified as using the a-versus-b rule. The low frequency of more sophisticated strategies is comparable to that seen in our prior research. Also, similar to our past results is the low rate of usage of the cell-a strategy. This is especially interesting, given that it is the most common of the proposed judgment strategies and was even said to be the modal strategy among adults (Smedslund, 1963; Nisbett & Ross, 1981). Our evidence finds this strategy to be rare among children as young as second grade.

These results would indicate that our prior procedures may have been unnecessarily confusing to younger subjects. Our prior and present procedures were not systematically compared in this paradigm, nor did we compare aspects of the changed procedure (e.g., instruction vs. question format) in a factorial design. As a result, we can offer little information about what aspects of the prior procedure may have been a problem. However, it is clear that we have developed a procedure suitable for use with young children. These findings indicate that children as young as second grade use simple but systematic rules in judging event relationships.

Age trends in this paradigm show origins of rule use in covariation judgment at age levels comparable to that of researchers in causal and probabilistic reasoning. However, in one respect, these results differ from Siegler's (1981) data on children's probability judgments. In those experiments, substantial numbers of children used the conditional probability rule by 8-9 years. In fact, a comparison of conditional probabilities was the modal response pattern in this age group in one of his experiments. In contrast, none of the subjects in this experiment was classified as using the conditional probability rule and only a few used the sum of diagonals rule. Our past research (Shaklee & Tucker, 1979; Shaklee & Mims, 1982) found the conditional probability rule to be used by only a minority of subjects even at adulthood. Thus, comparability between these paradigms in terms of early rule use is not matched by performance similarity in the later years. Expressing a judgment in terms of marbles in piles elicits more advanced rule use than a question asking for a comparable decision in terms of covariations between potentially related events. One difference may be that our problems are set in contexts of events that are readily interpreted as causally related. Adi and colleagues (1978) found that subjects used simpler, less accurate rules in evaluating cause-effect relationships than in making covariation judgments on analogous problems. Evidence such as this may indicate that covariation judgment in a causal context lags behind the same judgment about non-causal relationships.

Our evidence of systematic rule use at an early age is intriguing, but equivalently interesting are the unsystematic judgments of so many age peers. That is, at second and third grades a majority of children are classified by one of our rules (59% and 61% respectively), but a substantial minority in each grade produce unsystematic judgment patterns (19% and 39% respectively) or pass-no problem type criteria (Strategy 0 = 22% of second graders). Inspection of individual subjects' judgment patterns failed to identify any alternative strategic bases of these responses. Thus some children are unsystematic in rule use at the same age as other children begin to show use of simple judgment strategies. What did these rule users know that allowed them to judge the problems in a systematic fashion? Several factors may differentiate these rule users from their unsystematic age peers.

One possibility may be that unsystematic subjects are not using the tabled frequencies at all, but rather are judging the event covariations on the basis of their prior expectations about the event relationships. For example, such children may decide that plants are more likely to be healthy when they get plant food based on their real world experience, regardless of the event frequencies in the problems they are asked to judge. Our instructions already caution subjects against making expectancy-based judgments but those instructions may be readily forgotten as the subject solves the problems.

Expectancy-based judgments may be a source of unclassifiable response patterns, but what leads others of these young subjects to adopt an a-versus-b rule? We suspected that the judgment question itself may direct children's attention to cells a and b of the contingency table. Asked if plants are more likely to be healthy when they get plant food or when they do not get plant food, a subject may look at these two event conjunctions (i.e., healthy plants-plant food, healthy plants-no plant food). A subjects must also attend to the comparative aspect of the question in order to employ the a-versus-b rule. Mastery of either the attention direction or comparative aspects of the judgment (or both) may be key competencies underlying the shift to a-versus-b rule use at these early ages.

These are plausible sources of development in covariation judgment, but their roles in the origins of systematic rule use have yet to be demonstrated. An approach often employed to model a naturally occurring developmental trend is a training paradigm. That is, one might identify a training program which teaches non-rule users the knowledge said to differentiate those subjects from rule-based age peers. Contents of a successful training procedure identify at least one sufficient model to account for the natural transition to systematic rule use.

Experiment 2

We propose to use this training strategy in Experiment 2 to investigate the origins of systematic rule use in judging event covariation. Results of Experiment 1 indicated that reliable rule use was already becoming common in the second grade sample. Thus, Experiment 2 was an attempt to train first and second grade children to use the a-versus-b rule. We chose not to train children in use of the cell-a rule since it so rarely occurred naturally.

If young children's judgments are unsystematic because they are expectation-based, this problem would best be treated by drawing children's attention to the frequency information in the tables. Thus, one training procedure directed children's attention to the frequencies involved in the a-versus-b rule, i.e., cells a and b. This was the reasoning behind the Attention-only condition, where, on a set of 6 training problems, the experimenter asked the subject to point to the event combinations specifically mentioned in the question and to count the number of cases in each of the two cells. Subjects then made their covariation judgment.

As suggested previously, a subject may also fail to use the a-versus-b rule because he or she misses the comparative aspect of the question i.e., which is more likely. A second group of subjects were given the Attention instructions on the training problems and, in addition, were specifically asked which of the two cells had more cases in it. Subjects then made their covariation judgments. This group is the Attention-plus-more training group.

A final group is a no-training control group, who judged the same 6 problems but were given no special instructions.

All subjects were pretested to establish initial rule use. Unclassifiable, Strategy 0 and cell-a judges were included in the paradigm. Subjects were randomly assigned to one of the three conditions. Training effects were measured in a posttest given about a week after the training session. In view of their comparability in Experiment 1, problem set 24 and set 36 were problems in this experiment.

Method

Subjects

Subjects were respondents to ads in a small town newspaper offering first and second graders payment for participation in a psychology experiment. Forty-nine subjects participated in the pretest session of the experiment. However, 13 subjects were dropped from the experiment because their pretest strategy indicated that they were already using the a-versus-b (9 subjects) or a more advanced strategy (3 sum-of-diagonals subjects, 1 conditional-probability subject). The remaining 36 subjects (18 males and 18 females) included 13 unclassifiable, 17 Strategy 0, and 6 cell-a subjects. Mean age of these subjects was 7 years-6 months (range 6 years-10 months to 8 years-0 months).

Pretest

Problems and instructions on the pretest were identical to those described in Experiment 1. Half of the subjects were given problem set 24 for the pretest and set 36 for posttest, the remaining subjects were given the problem sets in the reverse sequence.

Once the problem set was completed, the experimenter determined the subject's judgment strategy in the manner described in Experiment 1.

Training

Six new problems were developed for training. These problems used cell frequencies and contents which were different from those used in the two test sets. Subjects classified as cell-a, Strategy 0, or unclassifiable were randomly assigned to one of three training conditions (12 subjects per condition).

Attention-Only. This training was designed to direct subject's attention to the two event pairings specifically mentioned in the question (i.e., cells a and b). Verbatim instructions for this condition were as follows (portions were re-phrased if necessary):

In doing these problems, you may have had a certain way of deciding which answer you thought was right. For example, you may have thought that certain boxes and the pictures in them were important and other boxes were not important in answering the question. Or you may have compared certain boxes with each other. If one thing happened more than another thing, it may have been more likely to happen. Now we are going to see if there might be another way to solve these problems that may be better than the way you used. We will try to decide which boxes and the pictures in them are important in deciding which answer is right. I want you to think hard now about a good way to answer these problems. I'll ask you some questions to help figure out a way to decide what answer is right. (The first problem and question were read to the child.)

If we wanted to decide which answer is right, it is important to look at each answer and find good examples or pictures that may show that thing happening. For example, let us suppose we wanted to see if answer A might be the right answer. Answer A says (e.g., the bugs are more likely to crawl on the leaves when it is sunny out). Could you show me which box or pictures are good examples of that? Which pictures show where the (bugs crawl on the leaves when it is sunny out)?

(Subjects should point to cell a, and were corrected if they did not. When subjects did point to cell a:)

Right. Can you tell me why? So these pictures show the (bugs crawling on the leaves when it is sunny out). This is an important box to look at in deciding if answer A is right. And how many times did that happen? So there are ______. good examples of answer A.

(The experimenter also pointed to other cells, asked or pointed out why they were not good examples.)

Now let us look at answer B, because that could also be the right answer.

(The same procedure was repeated. Subjects should point to cell b. The experimenter selected answer A and answer B to be discussed first with approximately equal frequencies. The discussion was then summarized.)

Okay, so that means that if we wanted to see if (question with answer A is read) this box (cell a) and the pictures in it would be important to look at. And we see that it happened _____ times. If we wanted to see if (question with answer B is read) this box (cell b) and the pictures in it would be important to look at. And we see that this happened _____ times. It is also possible that answer C is correct, that it didn't make any difference (if it was sunny or not, the bugs were just as likely to crawl on the leaves).

The covariation judgment question was then read to the subject and he or she made a response.

Attention-plus-more. This training condition was designed to emphasize the comparative aspect of the question, i.e., which outcome was more likely? The training builds on the Attention-only training described earlier. Subjects in this condition heard all of the instructions in the Attention-only training, and were then asked to make a direct comparison of cell a and cell b frequencies ("Which of these two things happened more?"). The experimenter then read the covariation judgment question to the subject and he or she made a response.

Control. Subjects in this condition judged the same problems as subjects in the other groups, but were offered no training instructions.

In each training condition, the procedure described was repeated on the six training problems. Feedback (positive or negative) was not provided following the subject's answers to the covariation judgment question.

Posttest

Subject fatigue prevented an immediate posttest of training effects. However, all subjects did return approximately one week later for a delayed posttest. This posttest was administered by a second experimenter who was blind to the training condition of the subject. The experimenter first reviewed the stimulus materials and problem format by presenting one of the sample problems used in session 1. Following this, the second problem set was administered in the same manner as in session 1. Subjects were tested on the problem set (24 or 36) not judged in the pretest session. Following completion of the problem set, subjects were told the purpose of the experiment and its potential relevance to everyday causal reasoning.

Results

The first indication of the relative success of the training methods was children's performance on the 6 training problems. Subjects responded in the manner predicted by the a-versus-b rule on 43.1% of the problems in the control group, 72.2% of the problems in the Attention-only group, and 97.2% of the problems in the Attention-plus-more group. An overall analysis of variance indicates these differences to be reliable, F (2,33) = 18.81, $p < .001$. Pairwise comparisons indicate that each training group is significantly different from each of the other groups (Duncan's multiple range test, $p < 05$).

Effects of the training procedure are most clearly assessed by comparison of the posttest performance of subjects in the training and control conditions. These effects will be analyzed both in terms of the accuracy of subjects on the various problem types and in terms of their posttest strategy classifications.

For each subject, posttest judgment accuracy was assessed in terms of the percentage of correct judgments for each of the 4 problem types. These data were analyzed in an analysis of variance including problem type (4 levels) and subject's training condition (3 levels) as factors. This analysis indicated a significant main effect of problem type, $F(3,99) = 17.22$, $p < .001$, and a significant interaction between problem type and training condition, $F(6,99) = 5.78$, $p < .001$. As the means indicate in Table 5, Attention-plus-more subjects were

---

Insert Table 5 here

---

substantially more accurate on cell-a and a-versus-b problems than on sum of diagonals and conditional probability problems. Attention-only and control subjects' performance were similarly poor across problem types. The main effect of training condition was not significant.

Pretest and posttest strategy classifications were compared for each subject to note training effects. Judgment was said to have improved if a subject was classified as using the a-versus-b, sum of diagonals, or conditional probability strategy at posttest. Table 6 indicates the frequencies of improvement

---

Insert Table 6 here

---

of subjects in each of the three training conditions. In all cases subjects who improved were categorized as using the a-versus-b strategy.

An overall χ^2 shows these training effects to be significantly different between conditions ($\chi^2 = 11.02$, df = 2, $p < .01$). As indicated in the table, rates of improvement were at similarly low levels (25%) in the control and Attention-only conditions compared with substantial rates of improvement (83%) among Attention-plus-more subjects.

Discussion

These results offer clear evidence of the differential effectiveness of our various training conditions. First, spontaneous improvement from test to retest was rare among subjects in the control condition. This would suggest that these young subject's problems were not simply lack of familiarity with the problems.

Improvement rates were equally low in the Attention-only condition. This null effect indicates that simply directing attention to cells a and b is not sufficient to elicit a-versus-b rule use among these children. The failure of Attention-only instructions may imply that subjects at this age already know how to find the cells mentioned in the question. If this were the case control and Attention-only subjects would be essentially equivalent in knowledge state at posttest. One would also expect that the Attention-only training would be sufficient to overcome any tendency to make expectation-based judgments. That is, children's attention was repeatedly directed to the information in the table cells. Indeed, the children's improved performance on the training problems suggests that the training was successful in eliciting frequency-based judgments. However, those effects were not maintained at the posttest one week later. Of course, any null effect has at least one alternative interpretation. That is, the Attention-only training condition may have simply been ineffective at teaching children the knowledge that should have been sufficient to elicit a-versus-b judgments.

However, the Attention-plus-more training did result in reliable improvement at the posttest. This finding indicates that the comparative aspect of the judgment may be a key obstacle to natural use of this simple rule by young subjects. Although they may know that two cells of the table are relevant, apparently subjects this young cannot spontaneously derive a way to combine that information to make a single judgment. Our training in the "more" rule apparently offers them that information. Since this training builds on the information offered in the Attention-only condition, this effect may hinge on the combined influence of the attention direction and comparative aspects of the question. Unfortunately a "More-only" condition is logically impossible. One cannot talk about comparing cells without designating which cells are to be compared. The fact that these training effects held over a one week delay period indicates the reliability of knowledge the children acquired.

Finally, it is worth noting the specificity of our training effects. That is, all children who improved in strategy use showed use of the a-versus-b strategy. This aspect of the results indicates that subjects were not simply learning to be systematic in judgment bases. Rather, they acquired one specific judgment rule. On this problem set, use of the a-versus-b rule did not lead to an overall improvement in judgment accuracy. This is by design of the problem set. That is, a-versus-b judges should be correct on cell-a and a-versus-b problems but incorrect on the sum of diagonals and conditional problems. Thus, the successful Attention-plus-more training actually results in worse performance of half of the problems compared to the other two conditions.

These training effects offer one sufficient model of the natural process of acquiring the a-versus-b rule. That is, subjects whose attention was directed to cells a and b and who were instructed to compare the two cells showed a-versus-b rule use. Thus, these two knowledge components may be the source of children's natural shifts to a-versus-b rule use. Of course, a sufficient process is not always a necessary one. That is, children may spontaneously discover the rule through yet another sufficient process.

These training effects may also be appreciated in a broader context. That is, research in causal reasoning indicates that some simple understanding of event covariation may begin in early elementary school (Shultz & Mendelson 1975; Siegler & Liebert, 1974). Siegler's (1981) work in probability judgment shows similar age trends in children's use of simple rules in comparing probabilities. This evidence indicates that those competencies may be shown at an even earlier age with a brief training procedure. It may be interesting to see if these improvements in covariation judgment would influence children's causal reasoning as well. This may be a domain in which to test children's ability to apply statistical concepts appropriately to related judgments.

Whether children could learn to use a more complex rule with appropriate training is a question for future research. However, the level of math involved in our other rules may preclude their use in early elementary school. The sum of diagonals rule requires a comparison of two sums, the conditional probability rule compares two ratios. These advanced arithmetic competencies are likely to be outside of the capacity of such young children.

In overview, these two studies offer new information about covariation judgment in the early elementary school years. That is, many children spontaneously show use of the a-versus-b rule as early as second grade. Children as young as first grade can be taught to use this simple rule if offered the relevant information. This training evidence offers one sufficient model of the natural acquisition of a simple rule for judging relationships between events.

References

Adi, H., Karplus, R., Lawson, S., & Pulos, R. Intellectual development beyond elementary school VI: Correlational reasoning. School Science and Mathematics, 1978, 78, 675-683.

Bullock, M., Gelman, R., & Baillargeon, R. The development of causal reasoning. In W. Friedman (Ed.), The Developmental Psychology of Time. New York: Academic Press, 1982.

DiVitto, B., & McArthur, L. Developmental differences in the use of distinctiveness, consensus, and consistency information in making causal attributions. Developmental Psychology, 1978, 14, 474-482.

Inhelder, B., & Piaget, J. The Growth of Logical Thinking from Childhood to Adolescence. New York: Basic Books, 1958.

Jenkins, H., & Ward, W. Judgment of contingency between responses and outcomes. Psychological Monographs, 1965, 79, 1-17.

Kelley, H. Attribution theory in social interaction. In E. Jones, et al., (Eds.), Attribution: Perceiving the causes of behavior. Morristown, NJ: General Learning Press, 1972.

Nisbett, R., & Ross, L. Human Inference: Strategies and Shortcomings of Social Judgment. Englewood Cliffs, NJ: Prentice-Hall, 1980.

Shaklee, H., & Mims, M. Development of rule use in judgments of covariation between events. Child Development, 1981, 52, 317-325.

Shultz, T., & Mendelson, R. The use of covariation as a principle of causal analysis. Child Development, 1975, 46, 394-399.

Siegler, R. Defining the locus of developmental difference in children's causal reasoning. Journal of Experimental Psychology, 1975, 20, 512-525.

Siegler, R. Three aspects of cognitive development. Cognitive Psychology, 1976, 8, 481-520.

Siegler, R. Developmental sequences within and between concepts. Monograph of the Society for Research in Child Development, 1981, 46, Whole No. 189.

Siegler, R., & Liebert, R. Effects of contiguity, regularity and age on children's causal inferences. Developmental Psychology, 1974, 10, 574-579.

Smedslund, J. The concept of correlation in adults. Scandinavian Journal of Psychology, 1963, 4, 165-173.

Footnote

[1]We had some difficulty defining a noncontingent relationship for the sum of diagonals problems. The problem we included (middle problem, column 3, Tables 2a and 2b) deviates slightly from independence ($P(A_1/B_1) - P(A_1/B_2) = -.06$, set 24, $-.03$ set 36) by the conditional probability rule. As a result we scored responses as correct if subjects concluded that A_1/B_1 was either less likely or just as likely as A_1/B_2. The problem does discriminate appropriately between the other judgment rules. Cell-a and a-versus-b judges should say that A_1/B_1 is more likely than A_1/B_2, sum of diagonal judges should say the two outcomes are equally likely.

Table 1

Contingency Table Cell Labels

	B_1	B_2
A_1	a	b
A_2	c	d

Table 2

A) Cell frequencies used for problems in problem set 24.

Cell a Problems

(11)	B_1	B_2
A_1	11	2
A_2	4	7

(3)	B_1	B_2
A_1	6	6
A_2	6	6

(5)	B_1	B_2
A_1	2	11
A_2	7	4

a versus b Problems

(6)	B_1	B_2
A_1	7	3
A_2	2	12

(9)	B_1	B_2
A_1	3	3
A_2	9	9

(4)	B_1	B_2
A_1	4	11
A_2	8	1

Sum of Diagonal Problems

(2)	B_1	B_2
A_1	2	2
A_2	2	18

(7)	B_1	B_2
A_1	9	7
A_2	5	3

(10)	B_1	B_2
A_1	8	8
A_2	8	0

Conditional Probability Problems

(8)	B_1	B_2
A_1	2	12
A_2	0	10

(12)	B_1	B_2
A_1	1	5
A_2	3	15

(1)	B_1	B_2
A_1	12	2
A_2	10	0

B) Cell frequencies used for problems in problem set 36.

Cell a Problems

(11)	B_1	B_2
A_1	16	4
A_2	6	10

(3)	B_1	B_2
A_1	9	9
A_2	9	9

(5)	B_1	B_2
A_1	3	16
A_2	10	7

a versus b Problems

(6)	B_1	B_2
A_1	11	2
A_2	7	16

(9)	B_1	B_2
A_1	5	5
A_2	13	13

(4)	B_1	B_2
A_1	6	18
A_2	9	3

Sum of Diagonal Problems

(2)	B_1	B_2
A_1	4	4
A_2	4	24

(7)	B_1	B_2
A_1	12	9
A_2	9	6

(10)	B_1	B_2
A_1	11	11
A_2	11	3

Conditional Probability Problems

(8)	B_1	B_2
A_1	3	18
A_2	0	15

(12)	B_1	B_2
A_1	2	7
A_2	6	21

(1)	B_1	B_2
A_1	18	3
A_2	15	0

Table 3

Strategy Classification Criteria

Strategy use and resultant patterns of problem accuracy.

(+ = accurate, 0 = inaccurate)

		Problem Strategy Type			
		Cell a	a versus b	Sum of Diagonals	Conditional Probability
Subject Strategy Type	Conditional Probability	+	+	+	+
	Sum of Diagonals	+	+	+	0
	a versus b	+	+	0	0
	Cell a	+	0	0	0
	Strategy 0	0	0	0	0

Table 4

Experiment 1

Rule classifications of subjects by Grade (percentages)

Grade	Strategy Classification						
	Unclassifiable	Strategy 0	Cell-a	a-versus-b	Sum of Diagonals	Conditional Probability	N
2	19	22	16	40	3	0	37
3	39	0	11	44	6	0	18
4	6	0	18	71	6	0	17

Table 5

Experiment 2

Mean percent correct for each problem type

Training condition	Problem Type Cell-a	a-versus-b	Sum of Diagonals	Conditional Probability	All
Attention-plus-more	83.3	80.6	8.3	5.5	44.4
Attention-only	55.4	44.3	27.8	33.3	40.2
Control	52.8	38.8	44.4	24.8	40.2
All	63.8	54.6	26.8	21.2	41.6

Table 6

Effects of a-versus-b training on posttest performance

	Improved	Didn't Improve	Total
Control	3	9	12
Attention Only	3	9	12
Attention plus more	10	2	12
Total	16	20	36