Four strategies used in judgment patterns were explored. Problem sets in which each solution strategy produces a unique solution pattern are depicted. Several experiments had been conducted using rules in this way with subjects from grade 4 through college. Problems were set in the context of concrete events which could be related, and subjects were asked about the relative likelihood of an outcome. A strong developmental trend was found, with students using increasingly sophisticated rules with increasing age. Current efforts are focused on trying to account for trends, noting knowledge differences between age groups that may be implicated in the differences in rule use. An experiment to train seven-year olds to use the $a$-versus-$b$ rule is described; this was successful. Next, students in grades 4, 5, 7, 8 were trained to use the sum-of-diagonals rule, again successfully. (MNS)
Training for Improved Covariation Judgment

Harriet Shaklee, Laurie Hall, and Don Paszek
University of Iowa


This research was supported by a National Institute of Education Grant G-80-0091 awarded to the first author. Many thanks to Paul Holt and Nancy Oetken for assistance in data collection.
A variety of theorists have suggested that covariation judgment may be a key element in causal reasoning. That is, people may find likely causes of an event by searching for covariates of that event. If causal and covariation judgment are interlinked in this way, then accuracy of covariation judgment may set an upper limit to an individual's competence at causal reasoning.

Evidence from our own investigations indicates that people show wide individual differences in competence at covariation judgment. In particular, a majority of adults employ rules which may lead to better than chance accuracy, but which result in systematic errors on same event relationships. We've focused our investigation on four strategies which might account for subjects' judgment patterns. Each of these strategies will be discussed in terms of the four cells of a 2 x 2 contingency table, labeled cells a, b, c, and d in a left to right, top to bottom sequence. One commonly proposed strategy is to judge a relationship according to the number of times the target event states co-occur, cell-a of the contingency table. We term this strategy the cell-a strategy.

A second approach might compare the number of times the target event occurs with its supposed cause with the number of times that event occurs without that possible cause. This strategy would compare frequencies in contingency table cells a and b, a strategy we call a-versus-b. A third strategy might compare the number of events confirming a relationship of target event and supposed cause with the number of events which would disconfirm such a relationship. This strategy would compare the sum of frequencies in cells a and d with that of cells b + c, a strategy we term sum of diagonals ((a + d) - b + c)). Finally, a mathematically sophisticated approach would compare the probability of target event
given the supposed cause with the probability of the event when that
cause was absent. We call this strategy the conditional probability
strategy and it is the only one of our strategies which will always produce
correct judgments of event covariations.

Thus, we propose four different judgment rules varying in complexity
and likely accuracy. Since different rules should produce different
judgments, we can construct a problem set where each solution strategy
produces a unique solution pattern. A sample of such problems is illustrated
in Table Ia. Problems are structured hierarchically such that cell-a
problems are accurately judged by all rules; a-versus-b problems should
be correctly judged by all but cell-a judges. Sum-of-diagonals problems
should be accurately judged by sum-of-diagonals and conditional-probability
problems should be accurately judged by the conditional probability
rule, alone. Accuracy of judgment is indexed by the direction of the
judged relationship. For example, a-versus-b judges should judge the
conditional probability problem in Table Ia as a case in which A1 is
less likely given B1, than given B2 (2-12). Sum of diagonals judges
should judge the two events as unrelated (2 + 10 = 0 + 12), and conditional
probability judges should see A1 as more likely given B1 than given B2
(2/2 vs. 12/22). A subjects' strategy is indexed by the accuracy pattern
on a 12 problem set, including 3 problems of each of the problem strategy
type. Table Ib indicates judgment accuracy predicted by each of the
proposed rules. Subjects who pass no problem types are labeled Strategy
0. All other patterns not represented in the table would be labeled
unclassifiable. We've looked at rule use in this way in several experiments
involving subjects from 4th grade through college age. Problems in
these experiments are set in the context of concrete events which could
be related. Frequency information is represented in pictorial format in a $2 \times 2$ table. Subjects are asked about the relative likelihood of an outcome given the two alternative states of the other variables.

Our past evidence indicates a strong developmental trend in the 4th grade to college age span. The modal strategy at 4th grade was the $a$-versus-$b$ rule, although Strategy 0 and unclassifiable judges were also common. The sum of diagonals rule was used by a substantial group of subjects in our 7th and 10th grade samples. The conditional probability rule was used by a substantial minority of subjects in tenth grade and college. The cell-$a$ rule was rare at all ages tested. Thus, subjects used increasingly sophisticated rules with increasing age. However, the optimal conditional probability rule was used by a minority of subjects even at college age.

Having discovered these developmental trends, our current efforts are trying to account for those trends. That is, what knowledge differences between these age groups may be implicated in the differences in rule use. A common approach to the problem is to develop a training method which is effective in eliciting use of more advanced rules. Contents of those effective interventions allow us to identify one sufficient account of naturally occurring developmental trends. Effective training programs may also be of pragmatic value in improving covariation judgment.

Our first concern was with the many fourth graders who didn't match any of our proposed rules. Given the number of such subjects, we have to consider the possibility that these children were confused by some aspect of our method and were unable to demonstrate their true competencies. Our approach was to elaborate our instructions to insure that the children understood the tabled stimuli and to reformulate the covariation question.
in a syntax more appropriate for younger children.

These modifications were made to make our problems more comprehensible to younger children. It turns out that we outdid ourselves in this respect. Testing a new sample of children, nearly all of our subjects were classifiable by one of our rules in the fourth grade, and a majority of children showed systematic rule use in the second and third grades. Overwhelmingly, these subjects were classified as using the a-versus-b rule. Unclassifiable and Strategy 0 judgment patterns were predominant among first and second grade children. As a result, this population was the target age for an attempt to elicit use of a simple judgment rule. Thus, the first experiment I'll describe is an attempt to train 7-year-old subjects to use the a-versus-b rule. We opted not to train children in use of the cell-a rule since it so rarely occurred naturally.

Our training approach stemmed from our suspicion that the judgment question itself focused children's attention on cells a and b of the contingency table. Asked if plants are more likely to be healthy when they get bug spray or when they don't get bug spray, a subject may look at those two event conjunctions (i.e. healthy plants-bug spray; healthy plants-no bug spray). We thought of this as a problem of attention direction. This was the reasoning behind our 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. Subjects had mastered this technique by the end of the training problems.

A subject may also fail to use the a-versus-b rule because he or she misses the comparison 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. Subjects also mastered this technique by the end of the training problems. 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.

Subject fatigue prevented an immediate posttest of training effects. However, all subjects did return a week later for a delayed posttest. Subject's performance at that time is illustrated in Table 2 of your handout. As you can see, rates of improvement were at the same low level for Attention-only and control subjects. This failure of Attention-only instructions may imply that subjects at this age already know how to find the relevant cells. However, the Attention-plus-More training did result in reliable improvement at the delayed posttest. Thus, we see that the comparative aspect of the judgment may be a key obstacle to natural use of this simple rule by young subjects.

Having discovered that young children could use this simple rule, we next attempted to elicit use of more advanced rules from older subjects. Our first approach was to train subjects to use the sum-of-diagonals strategy. This strategy is built on the notion that some event combinations confirm a particular relationship between events and that some combinations disconfirm that rule. For example, if bug spray is good for plants, we should see many cases of healthy plants with bug spray and unhealthy plants without bug spray. Healthy plants without bug spray and unhealthy
plants with bug spray would be exceptions to the relationship. Sum-of-
diagonals training taught subjects that cells a + d were good examples
of a positive relationship and that cells b + c were exceptions to the
rule. Subjects learned that the reverse was true for negative relationships.
Subjects practiced pointing to the cells with good examples and those
with exceptions to the rule on each of 6 training problems. Subjects
also counted the number of cases in cells a + b and in cells b + c for
the training problems. These subjects then made their covariation
judgments. A group of control subjects made covariation judgments on
the same problems without the benefit of training. Training effects
were measured in an immediate posttest and in a delayed test one week
later. Subjects in the experiment were 4th, 5th, 7th and 8th grade
children whose pretest performance showed use of cell-a and a-versus-b
rules.

The results of this training experiment are shown in Table 3. Note
that unclassifiable posttest subjects were not included in the analyses.
Trained subjects were significantly more likely to show use of the sum-
of-diagonals rule both at the immediate and at the delayed posttest.
This evidence indicates that subjects can indeed show improved rule use
with a relatively simple training procedure. These training procedures
were similarly effective among the younger and older subjects in the
sample. Our training in confirming and disconfirming cases not only
yielded better accuracy, but those judgments also conformed to the
pattern predicted by the sum-of-diagonals rule. This suggests that this
reasoning may well underly the natural acquisition of this rule in
children's development. At a minimum, these training effects identify
one sufficient model of this developmental process.
Our final efforts at training are looking at what it takes to elicit use of the optimal conditional probability rule among junior high aged subjects. Thus far it looks like our training efforts are successful. This set of training studies suggests that subjects at all ages may show problems in covariation judgment but that those problems are not irremediable. Our evidence suggests that relatively simple training efforts can elicit use of more sophisticated and more accurate judgment rules.
Training for Improved Covariation Judgment

Harriet Shaklee, Laurie Hall, and Don Paszek

University of Iowa


Table 1

A) Sample covariation problems

<table>
<thead>
<tr>
<th>Cell a</th>
<th>a versus b</th>
<th>Sum of Diagonal</th>
<th>Conditional Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>B2</td>
<td>B1</td>
<td>B2</td>
</tr>
<tr>
<td>A1</td>
<td>11</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>A2</td>
<td>3</td>
<td>16</td>
<td>15</td>
</tr>
</tbody>
</table>

B) Strategy use and resultant patterns of problem accuracy.
(± = accurate, 0 = inaccurate)

<table>
<thead>
<tr>
<th>Problem Strategy Type</th>
<th>Cell a</th>
<th>a versus b</th>
<th>Sum of Diagonals</th>
<th>Conditional Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>±</td>
<td>±</td>
<td>±</td>
<td>±</td>
</tr>
<tr>
<td>Strategy Type</td>
<td>±</td>
<td>±</td>
<td>±</td>
<td>±</td>
</tr>
<tr>
<td>Cell a</td>
<td>±</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Strategy 0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

This research was supported by a National Institute of Education Grant G-80-0091 awarded to the first author. Many thanks to Paul Holt and Nancy Oetken for assistance in data collection.
Table 2
Effects of a-versus-b Training on Delayed Posttest performance of 7 year old children

<table>
<thead>
<tr>
<th></th>
<th>Improved</th>
<th>Didn't Improve</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>3</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Attention Only</td>
<td>3</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Attention plus more</td>
<td>10</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>20</td>
<td>36</td>
</tr>
</tbody>
</table>

\[ \chi^2 = 11.02, \text{ df } = 2, \ p < .01. \]
Table 3
Effects of Sum-of-Diagonals Training on Immediate and Delayed Posttest performance on 4th-8th grade children

<table>
<thead>
<tr>
<th></th>
<th>Improved</th>
<th>Didn't Improve</th>
<th>Unclassifiable</th>
<th>Improved</th>
<th>Didn't Improve</th>
<th>Unclassifiable</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>4</td>
<td>17</td>
<td>2</td>
<td>5</td>
<td>14</td>
<td>4</td>
<td>23</td>
</tr>
<tr>
<td>Training</td>
<td>15</td>
<td>6</td>
<td>8</td>
<td>21</td>
<td>6</td>
<td>2</td>
<td>29</td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>23</td>
<td>10</td>
<td>26</td>
<td>20</td>
<td>6</td>
<td>52</td>
</tr>
</tbody>
</table>

\(\chi^2 = 9.6, \text{ df } = 1, \ p < .01\)

\(\chi^2 = 9.87, \text{ df } = 1, \ p < .01\)