DOCUMENT RESUME

ED 350 149 SE 052 771

AUTHOR Bullock, Merry
TITLE Scientific Reasoning in Elementary School: Developmental and Individual Differences.
INSTITUTION Max-Planck-Inst. for Psychological Research, Munich (West Germany).
PUB DATE Apr 91
NOTE 18p.; Talk presented at the Symposium on Scientific Thinking, SRCD (Seattle, WA, December 1991).
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Elementary Education; *Elementary School Science; Foreign Countries; Science Education; Science Instruction; *Scientific Methodology
IDENTIFIERS *Science Process Skills; *Scientific Thinking

ABSTRACT

Pre-adolescent children are generally characterized as incapable of applying scientific reasoning to test a causal relation. This paper describes research on children's scientific reasoning which shows that pre-adolescent children do have some systematic scientific reasoning skills. The subjects of this study were 260 second through fourth grade students and 34 adults in Munich, Germany. In the first part of the study, "Hypothesis Testing," students were asked it they can generate and recognize adequate experimental test strategies, and whether they can adopt hypothetical stance to predict outcomes on the basis of a hypothesized causal relation. In the second part, "Interpreting Evidence," children were shown information indicating whether a particular dimension was related to an outcome, and were asked to judge the causal relation and to justify their judgement. From the results, it is concluded that children do understand some of the requirements of an experimental test, at least by third grade. Specifically, they know that one must vary the dimension of interest. By fourth grade, children also understand that one must control other variable dimensions. The conclusion drawn from the second part of the experiment is that by third grade children can propose a contrastive empirical test and can accurately use information as evidence about a causal relation. By fourth grade they can adopt a "hypothetical" perspective to discuss how an outcome will vary if a potential cause is or is not relevant. (PR)

MAX-PLANCK-INSTITUT FÜR PSYCHOLOGISCHE FORSCHUNG

8000 München 40, Leopoldstrasse 24

Scientific reasoning in elementary school: Developmental and individual differences

Merry Bullock

Talk presented at Symposium on Scientific Thinking, SRCD, Seattle 1991

Paper 12/1991

MAX-PLANCK-INSTITUTE FOR PSYCHOLOGICAL RESEARCH

Leopoldstrasse 24, D-8000 Munich 40, F.R.G.

Scientific reasoning in elementary school:

Developmental and individual differences

Merry Bullock

Presumably, each of us is an expert in scientific reasoning, if not by natural inclination, at least from years of training. What does this expertise entail? As already mentioned in the previous two papers (Amsel & Flach, 1991; Koslowski, Susman & Serling, 1991), two of the central features of expertise in scientific reasoning are an ability to construct a valid test to see whether one event is related to another and an ability to evaluate hypotheses about events on the basis of evidence, not prior expectations or beliefs.

How and when do such abilities develop? Traditionally, (e.g., Kuhn, Amsel & McLoughlin, 1988; Inhelder & Piaget, 1958), the answer has been "not before adolescence." Despite a robust ability to *detect* causal relations, pre-adolescent children are generally characterized as incapable of applying this ability in scientific reasoning tasks where they must systematically *test* a causal relation. Grade school children are characterized as having several problems:

First, they are said to lack a "hypothetical" perspective that would allow them to separate questions that ask whether something affects an outcome, that is, hypothesis-based questions, from questions that ask how to make an outcome occur, that is, pragmatic concerns.

Second, they are said to lack adequate strategies for testing potential causes, preferring to make confirmatory tests where only a potential cause is present, rather than contrastive tests where a potential cause is both present and absent; and

Third, children are said to lack an ability to properly interpret or use evidence to make a causal judgment, especially when the evidence contradicts their prior expectations or beliefs.

The purpose of this paper is to describe research that asks whether these deficits adequately describe grade school children's performance. I hope to accomplish two goals. The first is descriptive: I will ask how children between the 2nd and 4th grades perform on tasks that tap two components of scientific reasoning: constructing an empirical test, and interpreting evidence, and how these components change in the gradeschool years. In doing this, I hope to convince you that pre-adolescent children do have some systematic scientific reasoning skills.

My second goal is concerned with identifying some of the sources of individual differences in improvement in children's scientific reasoning skills. To do this, I will ask whether and how these skills are related to performance in other areas postulated to be related to scientific thinking, for example, logical reasoning and pre-formal operational skills such as combinations or detecting indeterminacy.

Design and Procedure

So, let me begin with the first goal, and describe the scientific reasoning task. As outlined in Table 1, it consisted of two parts, addressed to different aspects of scientific reasoning skills. In

the first part, *Hypothesis Testing*, we asked whether children can (1) generate and (2) recognize adequate experimental test strategies, and (3) whether they can adopt a hypothetical stance to predict outcomes on the basis of a hypothesized causal relation. In the second part, *Interpreting Evidence*, children were shown information indicating whether a particular dimension was related to an outcome, and were asked to judge the causal relation and to justify their judgment.

Table 1:

Hypothesis Testing

1) Can children spontaneously suggest an appropriate test when asked to see if a variable affects an outcome;
2) When provided a range of test objects, can children choose those that would provide an appropriate experimental comparison;
3) Can children contrast hypothetical outcomes for cases where a variable does / does have an effect?

Interpreting Evidence

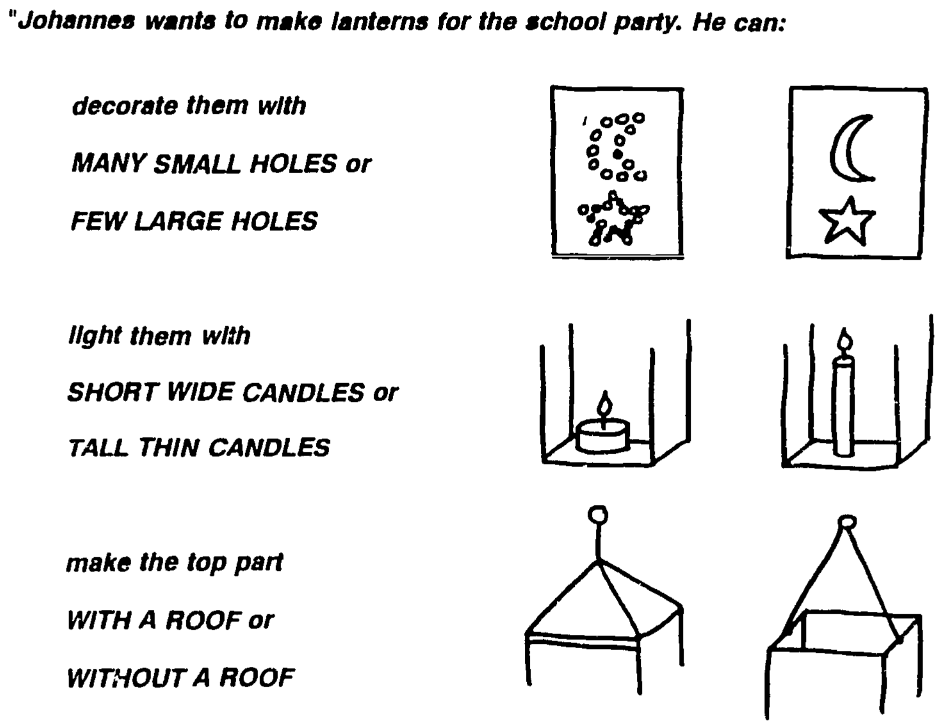
1) Can children use covariation evidence to judge whether a causal dimension is related to an outcome and to justify the judgment?
2) Is accuracy affected by prior expectations?

The subjects included 260 2nd through 4th graders and 34 adults. Of the children, 194 were part of an ongoing longitudinal study begun in Munich 6 years ago. Subjects received story problems in which a protagonist wanted to make some product, and wanted to test whether a particular dimension was important for producing successful outcomes. To make the task a little more real, I will describe the procedure

and results using one of the stories tested, a story about making lanterns.

The story, presented as a problem solving situation, was introduced with a series of pictures and the following text: "Johannes wants to make lanterns for his school party. He thinks it might be windy and wants to make sure the lanterns don't go out in the wind. He thinks about how to make the lanterns.... He can .. decorate them with many small holes or few large holes; light them with short wide candles or with tall thin candles; make the top part with a roof or without a roof .." (see Figure 1).

Figure 1

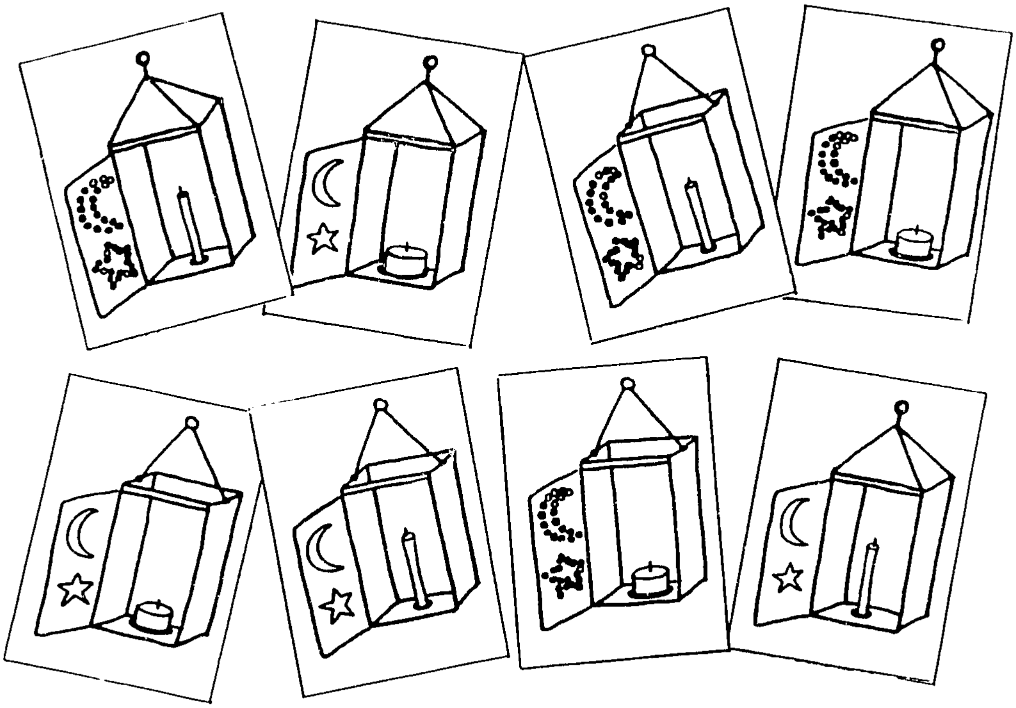


The text continued: "First, though, Johannes wants to find out whether how he makes the top part of the lantern makes a difference in how well a lantern will burn in the wind. What should he do to find this out?"

The first hypothesis testing measure was children's spontaneous *Verbal recommendations* to the question of how the protagonist should proceed. The second hypothesis testing measure was responses to a *Card choice* task illustrated in Figure 2, in which children were asked to choose a set of objects that would provide a critical test.

Figure 2

"Here are pictures of the lanterns .... Johannes could make. Which should he make to see whether the top of the lantern makes a difference?"



Answers to the *Verbal Recommendations* and *Card Choice* measures were coded in terms of whether subjects suggested varying the focal dimension (the roof of the lantern) and whether they held the other dimensions (holes and candles) constant. The precise categories were the following, as listed in Table 2:

Table 2

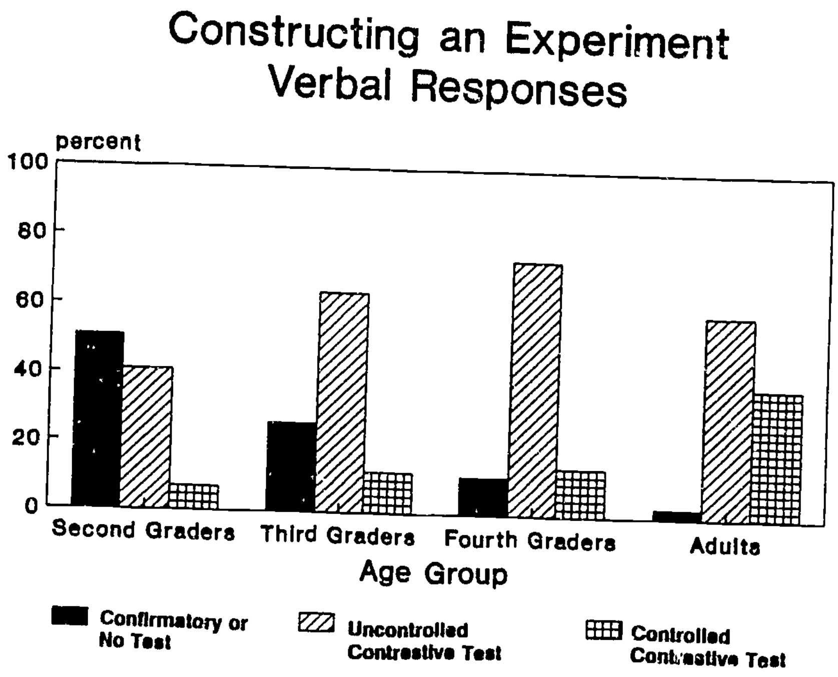
Measures	Coding
Verbal Responses Measure *Card Choice Measure*	Controlled Contrastive Test (focal dimension varied, others held constant: "make two just the same except one has a roof and one does not") Non-controlled Contrastive Test (only focal dimension varied: "make one lantern with a roof and a big candle, another without a roof and a little candle") Noncontrastive Test or No Test (focal dimension not varied; make just one lantern; no test needed)
Contrasting Outcomes Measure	Correct pattern: outcomes will vary if the focal dimension does matter, and will not vary if the focal dimension does not matter

For the third hypothesis testing measure, an ability to adopt a *hypothetical stance*, subjects were asked to contrast how outcomes would vary if the focal dimension did or did not make a difference. Responses to this question were coded as correct, as noted at the bottom of Table 2, or incorrect.

Results

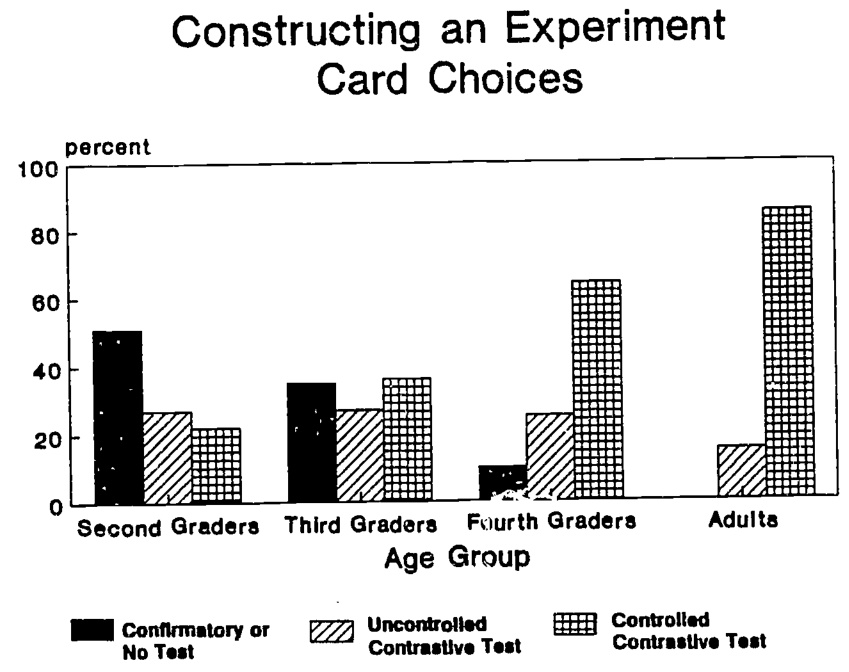
Figure 3 shows the distribution of subjects' verbal responses. As one can see, 2nd graders were mixed in this task: about half of them suggested a contrastive test. The other half proposed confirmatory tests, making just one lantern, or only lanterns of one type; in contrast, the large majority of 3rd and 4th graders (74%, 84%) proposed varying the focal dimension, a contrastive test. A small percentage of 3rd and 4th graders also added that one must hold one or both of the other dimensions constant, a performance that was not all that much worse than in adults.

Figure 3



The results from the card choice task, coded into the same categories, show the same general age patterns as those for the verbal responses, as seen in Figure 4. However, they also indicate that the performance level for the 3rd and 4th graders and adults was much higher: a third of the 3rd graders, most of the 4th graders, and almost all the adults could recognize or choose a critical test, even when they did not spontaneously suggest it. That is, although only 10 to 15% spontaneously said the protagonist should hold everything except the focal dimension constant, many more chose cards that did just that.

Figure 4



The validity of the card choice responses for showing competence is underscored by two sources of additional information: responses to the predicting outcomes task, and explanations of why children chose the cards they did. First, about 70% of those children who picked a critical comparison also correctly said that the two lanterns would burn differently if the roof did make a difference, and the same if it did not. Second, half of the children who picked a critical comparison in the card choice task, also explicitly justified this in terms of controlling dimensions.

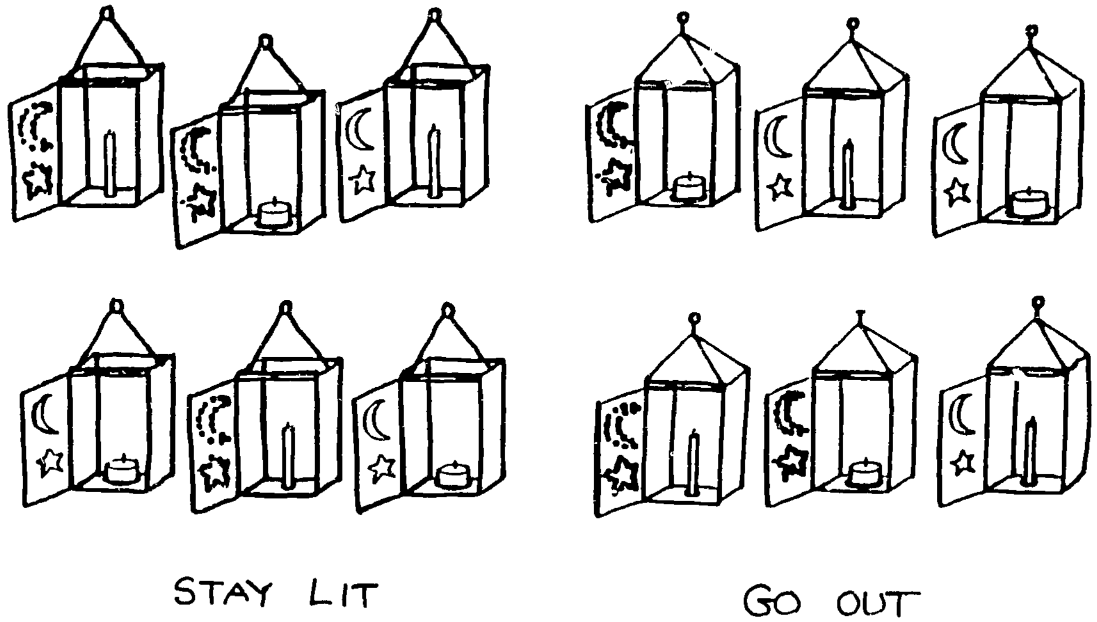
To conclude from this first part of the task, children do understand some of the requirements for an experimental test, at least by 3rd grade. Specifically, they know that one must vary the dimension of interest. By 4th grade, children also understand that one must control other variable dimensions, although this understanding is not yet reflected in their spontaneous suggestions for how to conduct an experiment. Thus concurring with Koslowski's (1991) conclusions, these data suggest that grade school children (at least by 4th grade) have a conceptual understanding of what an experimental test entails.

Let me now turn to the second part of the experimental task, the information interpretation part. Subjects were told that the protagonist had made several test objects, and had tried them out. They were then shown the outcomes and were asked whether the focal dimension was important or not, as illustrated in Figure 5.

We designed the information in this part so that it would be likely to contradict children's expectations, here that lanterns with

a roof would be better than those without. If a child did not interpret this information correctly, that is, if they said that the roof did not matter, they were shown a second set of simplified pictures with all dimensions except for the focal dimension held constant.

Figure 5

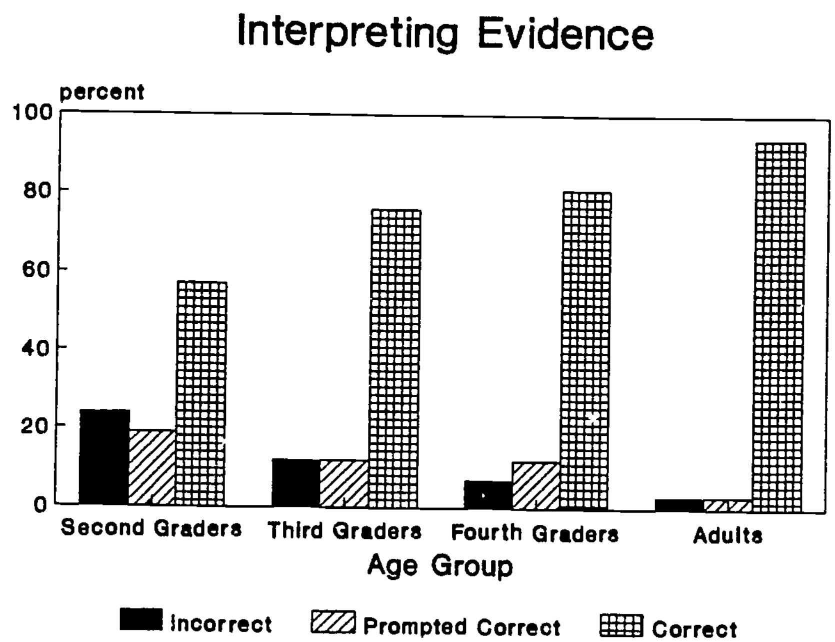


Answers were scored as correct, correct with prompt (the second picture) or incorrect. The results were fairly straight-forward: children made few errors on this part, as seen in Figure 6.

Children of all ages were able to interpret the simple covariation information, saying in our example that "no roof" was necessary for a good outcome; and there was a steady decrease in errors over age. Not only did children accurately interpret the information, they also justified their judgments by referring to the

evidence. Sixty-three per cent of the 2nd graders, 81% of the 3rd and 4th graders, and all of the adults justified their choices on the basis of the information about good and bad outcomes. Of these, almost half of the 2nd graders, 71% of the 3rd, 80% of the 4th and 91% of the adults specifically referred to the covariation of the focal variable with good and bad outcomes.

Figure 6



One the one hand, these results are not surprising: other studies have shown that children can generally interpret simple covariation information. On the other hand, because other studies have also added that children are not accurate when the information is <u>inconsistent</u> with their own expectations, these results are indeed

surprising, because we explicitly designed the information so that it would contradict children's expectations about what was important for a successful outcome. Of course, it might be argued that although we had designed the information to be inconsistent with children's expectations, we were simply unsuccessful. However, over the course of the procedure, it was possible to gather additional information about children's expectations about the focal dimension, either from their spontaneous utterances about what did and did not matter, or from our own probes.

We could thus look at the information interpretation data to see whether prior expectations made a difference. Children were coded into the three categories listed in Table 3, depending on whether or how the child would have to change his or her prior opinion to correctly interpret the information.

Table 3

No change -- the child expects the level of the focal dimension that really is associated with good outcomes to be associated with them (e.g., lanterns without a roof will burn better)

Change dimension -- the child expects the focal dimension to be irrelevant to good outcomes, when in fact it is relevant (e.g., the roof type doesn't matter)

Change level -- the child expects the focal dimension to be relevant but thinks the wrong level is associated with good outcomes (e.g., lanterns only burn well when they have a roof).

We then simply asked whether errors were disproportionately distributed among expectation types. The answer is basically "no", although there was a slight tendency for one type of conflict with prior belief to affect accuracy. When children had to change a belief that a dimension did not make a difference they were as accurate as those children whose beliefs were confirmed by the information. However, when children had to change which particular level was related to a positive outcome (e.g., from believing that a roof was important to seeing that no roof was important for a good outcome), they were more likely to err. It should be noted, however, that the majority of children with inconsistent prior beliefs were still accurate in interpreting the information.

To summarize: by 3rd grade, school children can in fact propose a contrastive empirical test and can accurately use information as evidence about a causal relation, usually even when it contradicts their expectations. By the 4th grade, they can moreover adopt a "hypothetical" perspective to discuss how an outcome will vary if a potential cause is or is not relevant, and are also somewhat aware of the need to not only vary but also to control variables.

Now I would like to briefly turn to my second goal, looking at individual differences. Even at the same age or grade level, there were substantial differences in how children performed, especially in the hypothesis testing part of the task. To look a little more closely at the sources of these differences, we were able to compare performance on the scientific reasoning task with measures of IQ, logical reasoning and pre-formal operational skills for the 194 longitudinal children. Because these children were given parallel

forms of the scientific reasoning task 1 year apart, we could ask which, if any of these skills were related to improvement.

We computed improvement measures as the difference between the same children's composite performance scores measured one year apart. Because some of the longitudinal children were in 2nd Grade at the beginning of the study, and some in 3rd grade, we conducted these analyses separately for the two groups of children. When I discuss the two groups of children, I will refer to those children who were in 2nd and 3rd Grades at the two measurement times as the younger children (mean age 8.6), and to those who were in 3rd and 4th Grades as the older children (mean age 9.1).

Among the younger children, improvement in the hypothesis testing part of the task was related to logical skills, whereas improvement was not related to any of the measures for the older children. What this suggests is that a minimal degree of logical skills are necessary for understanding how to construct an experiment, and that this minimal competence becomes available by 3rd Grade. Similar analyses for performance at each measurement point showed that logical skills were related to performance at Time 1 for the younger children, but not at Time 2 when they were in 3rd Grade, and logical skills were not related to the performance of the older children at either time. For both groups, performance level was additionally related to IQ.

In contrast to the hypothesis testing part where improvement for the younger, but not older children could be predicted on the basis of logical skills, improvement on the evidence interpretation and justification part was related to pre-formal operational skills, but

only for the older children. Performance at Time 1 was related to pre-formal operational skills for the older children, and at Time 2 for both groups of children. What this means is that some pre-formal operational skills (such as combinations, and detecting indeterminacy) may underlie the ability not only to reason about, but also to explicitly justify judgements about how information does or does not support a particular causal conclusion.

In conclusion, these data contribute to what seems to be, at least in this symposium (Amsel & Flach, 1991; Koslowski, et al., 1991; Sodian & Zaitchik, 1991), mounting evidence that grade school children's scientific reasoning skills are better and more systematic than their previous reputation would lead us to believe.

References

Amsel, E. & Flach, S. (1991). Developmental changes in evaluating evidence. Paper presented in the Symposium "Evidence Evaluation, conceptual change and the development of scientific reasoning", Society for Research in Child Development, Seattle.

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

Koslowski, B., Susman, A. & Serling, J. (1991) Conceptual vs. technical understanding of evidence in scientific reasoning. Paper presented in the Symposium "Evidence Evaluation, conceptual change and the development of scientific reasoning", Society for Research in Child Development, Seattle.

Kuhn, D., Amsel, E. & McLoughlin, M. (1988). The development of scientific thinking skills. New York: Academic Press

Sodian, B. & Zaitchik, D. (1991). Young children's differentiation of hypothetical beliefs from evidence. Paper presented in the Symposium "Evidence Evaluation, conceptual change and the development of scientific reasoning", Society for Research in Child Development, Seattle.

ERIC/EECE
University of Illinois
805 W. Pennsylvania
Urbana, IL 61801

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Scientific reasoning in elementary school: Developmental and individual differences	
Author(s): Bullock, Merry	
Corporate Source:	Publication Date: April 1991

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

	Sample sticker to be affixed to document	Sample sticker to be affixed to document	X
Check here Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction	"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."	"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."	**or here** Permitting reproduction in other than paper copy.
	Level 1	Level 2	

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature: Merry Bullock	Position: Senior Research Scientist
Printed Name: Merry Bullock	Organization: Max Planck Institute for Psychological Research
Address: Max-Planck Institute for Psychological Research Leopoldstr. 24 ~~8000 Munich / FRG~~	Telephone Number: (089) 38 602 222
	Date: December 20, 1991

SRCD 12/91