The purpose of this research was to explore concepts about thinking that contribute to children's metacognitive experience during experimentation. Twenty-four randomly selected elementary school children in Toronto, Ontario, Canada, participated in five theory of mind tasks that assessed their understanding of instantiation inference; understanding of causal and non-causal inference; understanding of causal and non-causal evidence; understanding of critical and non-critical evidence in a hypothesis testing format; and understanding of critical and non-critical evidence in a referential communication task. It was concluded that: children progress from unplanned to planned experimentation; the relationship between plans and procedures changes with grade level, so that early procedures do not depend on plans, but later procedures do; planning contrastive experiments does not depend on understanding causal inference; children's ability to direct and explain their own reasoning in science experimentation tasks was at the same level as their ability to affect and explain the beliefs of others in theory of mind tasks; and theory of mind appears to operate in two aspects of experimentation. In the productive phases it guides procedures by supporting the child's planning while in the critical phases it facilitates acceptance of disconfirming evidence and allows children to explain how their observations justify a particular inference.

(JRH)
The Contribution of Children's Understanding of Sources of Knowledge To Their Science Experimentation

Paper presented at the Symposium on
"Theory of Mind at School: Children's Scientific Reasoning"
At the annual meeting of the American Educational Research Association
San Francisco, California, April, 1995

This paper is a preliminary, partial report of research undertaken by the author to satisfy requirements for the PhD degree in Educational Psychology. The author wishes to acknowledge gratefully David R. Olson for his supervision of this research, and Keith Stanovich for his comments on an earlier presentation of this material.
The previous two presenters at this symposium focused on children's understanding of evidence as an aspect of their understanding of mental states. My research also concerned the link between children's theory of mind and their scientific reasoning. However, my concern was with the way in which the knowledge implicit in children's theory of mind guides their plans and actions as they actually produce evidence through science experimentation. Bringing together these two areas of research, children's theory of mind and the psychology of science experimentation, seems to offer benefits from the perspective of both: From the point of view of research on children's theory of mind, it may be asked if the declarative knowledge that this domain embodies might support the procedural knowledge that the child expresses during experimentation. From the viewpoint of the psychology of science education, researchers such as H. L. Swanson (1990) have already shown that metacognition significantly affects children's experimentation: the questions which theory of mind research might help to answer is, just which concepts about thinking contribute to children's metacognitive experience during experimentation? And what other, earlier-emerging, concepts might be developmental precursors to these?

In order to link these two areas, science experimentation was treated as a case of planning. A plan can be thought of as a mental representation of an action, or a series of actions, directed toward a goal. The actions are causes, and the goal is an effect of these actions. In order to form a plan, the individual
must have a mental model of the situation in question; she can then begin with the goal, and search backward through the model to find a cause. She can then implement this cause as an action in order to achieve the goal.

A science experiment which tests a hypothesis about a causal relationship is a specific sort of plan: an epistemic plan, that is, a plan to find something out. The goal is to know about a causal relationship—either that it obtains, or that it does not. In order to form a plan to test a hypothesis, the student needs a model of what comprises a cause of knowing about causal relationships. To put this another way, the student needs to know the source of causal beliefs, i.e. covariation evidence. This understanding of the source of causal beliefs is an aspect of the learner's theory of mind. This view of the role of theory of mind in experimentation implies several specific hypotheses which were tested in our research. Because time today is limited, research on three of these hypotheses will be described briefly, while results concerning the other three hypotheses will simply be reported.

**Method**

The research subjects were 24 children in each of grades one, three and five, randomly selected from three elementary schools in the Toronto area, to comprise a sample that was ethnically and socioeconomically diverse.

Each child participated in five theory of mind tasks: these assessed their understanding of instantiation inference (Sodian &
Wimmer, 1987); understanding of causal and noncausal inference (adapted from Ruffman, Perner, Olson & Doherty, 1993); understanding of causal and noncausal evidence; understanding of critical (unambiguous) and noncritical (ambiguous) evidence in a hypothesis testing format (Sodian, Zaitchik & Carey, 1991); and understanding of critical and noncritical evidence in a referential communication task.

Only the understanding of causal inference task (adapted from Perner et al., 1993), will be described here in detail. The materials included two pictures of children with healthy teeth, and two with unhealthy teeth, and pieces of red and green "food" which could be placed next to each. First, the researcher presented a picture paired with a piece of food, and said, "This boy ate some red food and he has healthy teeth. This girl ate some red food and she has healthy teeth. This boy ate some green food and his teeth are falling out. This girl ate some green food and her teeth are falling out." Then the researcher asked, "In this story, does one kind of food make children's teeth fall out? Which kind?" This question required the children to make their own causal inference.

Then the researcher reversed the relationship between the colour of food and the health of the teeth, and brought in a new character in the person of a doll, and asked, "Alison wasn't here before was she? When she sees the way things are now, will she think that one kind of food makes the children's teeth fall out? What kind?" In this operationalization, if the child represented Alison's inference, he or she said that the doll would think that
red food makes the children's teeth fall out. If the child could not represent Alison's inference, he or she typically attributed their own belief to the doll, saying that Alison would think green food makes the children's teeth fall out. Finally, in order to assess the child's understanding of noncausal inference, the researcher rearranged the evidence so there was no relationship between the colour of the food and the health of the character's teeth, then introduced a second doll, and repeated the questions. The order in which the causal and noncausal questions were presented was counterbalanced.

The children also participated in two Piagetian science tasks: In the pendulum task, children experimented to discover the factors which affect the speed of a pendulum. Typical hypotheses included the length of the string, the weight of the bob, the displacement, and the initial speed of the bob. In the "falling bodies on an inclined plane" (ramps and marbles) task, children rolled a marble down a ramp and off a small springboard, and tried to identify the factors that affected the distance that it jumped. Children could vary the height of the ramp, the size of the marble, the distance from which they released the marble, and the speed at which the marble started.

Results

Hypothesis 1: Children progress from unplanned to planned experimentation.

Table 1 represents the plans that children stated for the majority
of trials on their science experiments at various grade levels; the results will be illustrated using the example of the question of whether the height of the ramp makes a difference in how far the marble jumps. The columns, from left to right, represent increased planning, and validity of plans. Five of the grade one children offered no plan at all regarding the height of the ramp. Two planned a noncontrastive test, in which they would try only one level of the factor; e.g., they said that they would try a high ramp only. Five offered ad hoc plans; e.g. they initially said they would set the ramp high, then after doing so, they stated they would also try the ramp "low." Seven offered contrastive plans, stating that they would try the ramp at two heights, and see which went farther, but these children also varied another factor at the same time, such as the size of the marble, introducing a confound factor. Five children stated controlled plans, that is, they said that they would change the height, and keep other things, such as the size of the marble, "the same." The plans of the grade one children were diverse, while grade three and grade five children tended to plan contrastive or controlled tests, showing a trend toward improvement with grade level, $r(72) = .48, p < .01$.

Hypothesis Two: The relationship between plans and procedures changes with grade level, so that early procedures do not depend on plans, but later procedures do. Table 2 shows the relationship of plans to procedures that
each child used on the majority of their trials during the science experiments. Each row represents one of the types of plans discussed earlier. Generally, the level of plans correlated with the level of procedures, $r (72) = .79, p < .001$. However, of the five children who offered no plan, four actually made a contrastive test, for example, they tested the effect of the height by trying both a high and a low ramp. The same is true of the children who offered noncontrastive and ad hoc plans. In these three kinds of cases, the procedures were actually more complete than the plans, so these children could be characterized as exploring spontaneously. The fourth and fifth rows represent children who planned to contrast two levels of the factor, i.e. they planned at least two trials in advance. Here, in most cases, the children's procedures directly reflected their plans.

Consistent with the hypothesis, the relationship between plans and procedures was closely associated with grade level: For most grade one children, their procedures exceeded their plans; in contrast, for most children in grades three and five, plans and procedures matched; a Kruskal-Wallis nonparametric analysis of variance showed that this relationship was significant, $\text{chi}^2 (2) = 6.01, p < .05$.

Does this increase in the explicitness of plans with grade level only reflect the older children's ability to articulate their procedural knowledge, or does explicit planning actually affect the children's procedures? Note that in Table 2, the last column indicates that none of the children used controlled procedures
unless they had stated a plan describing at least two trials in advance. The implication seems to be that children can progress to the level of contrastive procedures without explicit planning, but in order to adopt controlled procedures, they must be able to articulate a plan.

**Hypothesis Three:** Planning contrastive experiments depends on understanding causal inference.

Recall that a contrastive experiment is one in which the child compared the effects of two levels of a factor. The hypothesis was that understanding causal inference would be necessary for planning a contrastive experiment. The first row of Table 3 indicates that of the students who did not predict someone else's inference, most could plan a contrastive test. The bottom row indicates that most children who understood causal inference could plan a contrastive test. This relationship remained significant when a loglinear test was used to control for effects of grade level and pretest knowledge, loglinear coefficient $^1$ = .29, standard error = .16, $Z = 1.78$, $p < .05$. These results suggest that, contrary to the hypothesis, understanding causal inference is not a necessary condition for planning an experiment, but that it may be a sufficient condition for doing so.

$^1$This coefficient is a parameter representing the relationship between the independent and the dependent variable, but the range is not limited to -1.0 to +1.0, and is not comparable to other indicators of covariation, such as the Pearson or Spearman $r$. 
Other Results

Because time is limited, I will not describe the way in which we tested each of the hypotheses that linked theory of mind to experimentation, but I will summarize them briefly.

The fourth hypothesis was that planning controlled experiments depends on understanding the notion of critical evidence. Children's ability to distinguish between critical and noncritical evidence was assessed using two tasks. One was Sodian, Zaitchik, and Carey's (1991) "mouse" task, which required the subject to select a way to unambiguously test whether an unseen mouse was large or small. The mouse task predicted the child's ability to design a controlled experiment, when pre-experimental physics knowledge and grade level were controlled statistically, loglinear coefficient = .32, standard error = .16, Z = 1.95, p < .05. The second task assessed the child's understanding of critical evidence by asking them to select between an ambiguous and an unambiguous message, and also significantly predicted children's ability to plan a controlled experiment, loglinear coefficient = .29, standard error = .15, Z = 1.86, p < .05.

The fifth hypothesis was that planning a controlled experiment requires an understanding of noncausal evidence. The rationale for this hypothesis was that control involves ruling out a possible causal relationship, and therefore relies on an understanding of noncausal evidence. The children's understanding of causal and noncausal evidence was assessed using the reverse of the healthy teeth task described earlier: the subjects were presented with a
causal belief different from their own, and asked them to arrange the evidence the story character "thinks she will see." As expected, identifying the evidence for an other's noncausal belief predicted controlled experimentation, \( r (71) = .37, p < .01 \); when grade level and pre-experimental physics knowledge were controlled statistically, this relationship was only marginally significant, coefficient = .23, standard error = .15, \( Z = 1.51, p = .1 \).

Previous researchers have found that when an individual holds a causal belief, and is confronted with noncausal evidence, they often retain their previous causal belief. The sixth hypothesis was that if children understood causal inference, they would be more likely to change their causal beliefs when they encountered disconfirming evidence. However, understanding of causal inference only marginally predicted changes in belief, \( r (62) = .22, p < .1 \).

The seventh hypothesis was that justifying a causal claim requires an understanding of causal evidence, and justifying a noncausal claim requires an understanding of noncausal evidence. At first, this may seem analytically true: to justify a causal claim simply is to understand causal evidence. However, recall that understanding of causal evidence was operationalized here as the subject's ability to state the covariation evidence for someone else's belief, while justification involved stating covariation evidence for their own beliefs. The relationship between these two measures was significant when pre-experimental physics knowledge and grade level were controlled statistically, coefficient = .33, standard error = .15, \( Z = 2.17, p < .05 \). An analogous relationship
was found for noncausal evidence, $r (51) = .31, p < .05$, however, when confounds were controlled statistically, this relationship was nonsignificant, coefficient = .16, standard error = .17, $Z = .91$, $p > .1$.

**Discussion**

The results of this research must be qualified by acknowledging that the findings are correlational, and the strength of most of the relationships, when potential confounds were controlled statistically, was moderate. However, we can draw some tentative conclusions. First, in most cases, children's ability to direct and explain their own reasoning in science experimentation tasks was at the same level as their ability to affect and explain the beliefs of others in theory of mind tasks. Second, theory of mind appears to operate in two aspects of experimentation: In the productive phases of experimentation, it guides procedures by supporting the child's planning. In the critical phases of experimentation, theory of mind may slightly facilitate acceptance of disconfirming evidence, and allows children to explain how their observations justify a particular inference.

Educationally, these results suggest the value of classroom discussions that encourage primary children to reflect on the sources of their scientific beliefs. One context for such discussion is the "discovery circle," in which the members of a class gather regularly during a unit of study to pose questions, show classmates their experiments, and talk about what they have
learned. Another context for reflective discussion is the "learning log," a diary in which children record what they have learned, and how they have learned it. It is hoped that by encouraging this kind of reflection, teachers will enhance children's theory of mind regarding causal reasoning, and thereby support children's ability to plan their own science experiments.
References


### Table 1

Children's Plans for Experimentation as a Function of Grade Level

<table>
<thead>
<tr>
<th>Grade</th>
<th>None</th>
<th>Noncontrast</th>
<th>Ad Hoc</th>
<th>Contrast</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Three</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>Five</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>4</td>
<td>9</td>
<td>20</td>
<td>34</td>
</tr>
</tbody>
</table>

\( r(72) = .48, p < .01 \)
Table 2

Level of Procedure as a Function of Level of Plan

<table>
<thead>
<tr>
<th>Plan</th>
<th>Noncontrastive</th>
<th>Contrastive</th>
<th>Controlled</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Noncontrastive</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Ad Hoc</td>
<td>0</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Contrastive</td>
<td>0</td>
<td>12</td>
<td>8</td>
</tr>
<tr>
<td>Controlled</td>
<td>0</td>
<td>2</td>
<td>32</td>
</tr>
</tbody>
</table>

r (72) = .79, p < .001
<table>
<thead>
<tr>
<th>Understanding of Causal Inference</th>
<th>Experimental Plan</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt; Less than</td>
<td>&gt; Contrastive</td>
<td>Or Controlled</td>
<td>Total</td>
</tr>
<tr>
<td>No</td>
<td>9</td>
<td>9</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>Yes</td>
<td>8</td>
<td>44</td>
<td></td>
<td>52</td>
</tr>
<tr>
<td>Total</td>
<td>17</td>
<td>53</td>
<td></td>
<td>70</td>
</tr>
</tbody>
</table>

$r (70) = .44, \ p < .01$