Scientific Reasoning Development: Children's Evaluation of Experiments and Evidence

Taneisha Vilma and Bradford H. Pillow

Northern Illinois University

Northern Illinois University

Background

We investigated scientific reasoning among elementary school students by assessing experiment and evidence evaluation abilities.

- The age of emergence for basic scientific reasoning skills is a continuing area of debate. Several researchers maintain that the fundamental skills involved in scientific reasoning do not develop until preadolescence (Kuhn et al., 1995, 2000; Schauble, 1996). However, others have argued that experiment and evidence evaluation skills are apparent during early childhood (Piekny & Maehler; 2013; Sodian et al., 1991).

- Previous studies have not fully distinguished these two abilities, or have assessed them with different tasks, making children's experiment evaluation and evidence evaluation performance difficult to compare.

The current study examined experiment evaluation and evidence evaluation within the same tasks to compare these two abilities directly.

We asked children to evaluate simple determinate and indeterminate experiments and their outcomes, and also asked children to evaluate experiments and evidence with two confounded or unconfounded variables.

Aims & Hypotheses

Aims

The first aim of the study was to compare (a) children's ability to evaluate whether an experiment is potentially conclusive or inconclusive and (b) children's ability to evaluate whether evidence is conclusive or inconclusive.

The second aim of the study was to investigate children's ability to recognize that the presence of confounded variables makes an experiment and evidence inconclusive.

Predictions

Experiment evaluation: For simple one variable experiments we expected that all age groups would perform above chance, but that performance would improve with age.

Evidence evaluation: For simple one variable problems we expected that performance would improve with age, but did not have prediction about the precise age at which children would be above chance.

Recognizing confounds: For two variable experiments, we expected that both experiment and evidence evaluation would improve with age for confounded experiments, but that all age groups would perform above chance for unconfounded experiments.

Method

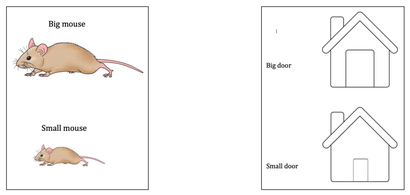
Participants

54 children elementary school children participated: 20 1st graders (M age = 7.0 years), 16 3rd graders (M age = 8.7 years), and 18 5th graders (M age = 10.8 years); 32 girls (60.4%) and 21 boys.

Procedure

Overview
All students participated individually. . An introduction story was provided followed by four control questions followed by the 1-Variable condition and the 2-Variable condition. Each trial included an experiment evaluation question followed by the evidence evaluation question.

At the beginning of the procedure, children were shown a picture of a box with a large opening and a box with a small opening. They were told that a small mouse can fit into either opening, but a large mouse can only fit into the large opening.



1-Variable Condition: 1-Variable condition examined children's ability to recognize the difference between conclusive vs. inconclusive information

The experimenter explained that two siblings disagreed about the size of a mouse in their home. One thinks it is a big mouse, but the other thinks it is a small mouse. The participants were asked to help the friends/siblings figure out the size of the mouse.

- **Inconclusive trial:** The siblings placed food in a box with a big opening. Participants were asked (a) "If Callie and Andy put the red cereal into the box with the big opening, will they know if the mouse is big or small?" and (b) "In the morning, the cereal is gone from the box with the big opening. Do they know if the mouse is big or small?"

- **Conclusive :** The siblings placed food in a box with a small opening. were told that siblings placed mouse food in the box with the small opening and asked (a) "If Sam and Riley put the red cereal into the box with the small opening, will they know if the mouse is big or small? and (b) "In the morning the food is gone, do they know if the mouse is big or small?"

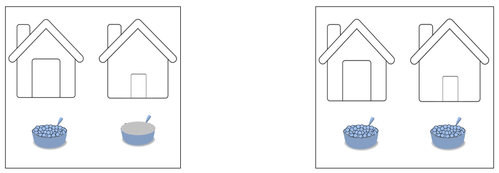
(a) *(b)*

2-Variable Condition: For the 2-Variable condition, participants were presented with two pictures of mouse houses, one with a large door and one with a small door and picture of how seven different color bags of cereal labeled 'mouse food'. The 2-Variable condition included two unconfounded trials and three confounded trials.

The experimenter explained that two friends disagreed about the size of a mouse at their school. One thinks it is a big mouse, but the other thinks it is a small mouse. The participants were asked to help the friends figure out the size of the mouse.

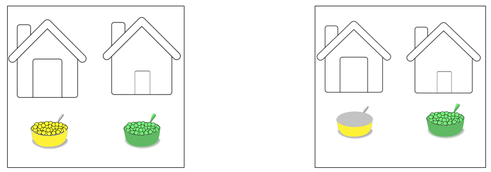
Unconfounded trial: (a) "Ryan and Liz put the blue cereal in both mouse houses. If the cereal is gone from the small house but not the big house, would they know what size the mouse is?" and (b) "In the morning, the food is gone from only the small house. Do they know for sure if the mouse were big or ?"

(a) (b)



- **Confounded trial:** (a) "Bailey and Alex put different types of food in the mouse houses. They put the yellow cereal in the big house and the green cereal in the small house. If the cereal is gone from the small house would they know if it is a big or a small mouse?" (b) "In the morning, the cereal is gone from the big house but not gone from the small house. Do they know for sure if the mouse is big or small?"

(a) *(b)*



Scoring
For the 1-Variable task, children received 1 point for each correct answer for each of the two questions. Scores for the experiment and evidence evaluation questions ranged from 0-3 correct. Scores for the conclusive question ranged from 0-4 correct and scores for the inconclusive question ranged from 0-2 correct.

For the 2-Variable task, scores were 0-5 for the experiment evaluation and evidence evaluation questions. Unconfounded evaluation score ranged from 0-4 correct and scores for the confounded evaluation ranged from 0-6 correct.

Results

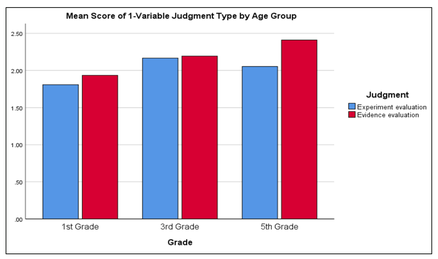
1-Variable condition: 3 x 3 x 2 (Age x Gender x Judgment) ANOVA did not reveal any significant effects.

Experiment evaluation: None of the age groups performed significantly different from chance for the experiment evaluation task.

- 1st grade: $t(19) = 1.61$, $p = .12$; 3rd grade: $t(15) = 2.07$, $p = .056$; 5th grade: $t(17) = 1.62$, $p = .12$.

Evidence evaluation: Children in all three grades performed significantly better than chance for the evidence evaluation task.

- 1st grade: $t(19) = 2.43$, $p = .02$; 3rd grade: $t(15) = 2.82$, $p = .01$; 5th grade: $t(17) = 6.86$, $p < .001$.

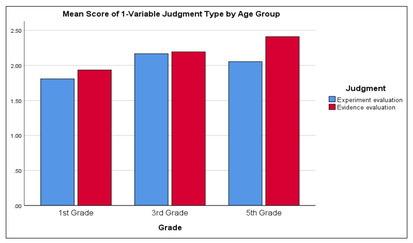


One sample t-tests compared performance on conclusive and inconclusive trials to chance (combining performance on experiment and evidence evaluation).

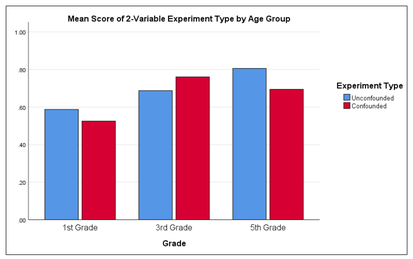
Conclusive trials: 1st grade: (M =2.90, SD =1.21); $t(19) = 3.32$, $p = .004$ and fifth grade: (M =3.05, SD =1.05); $t(17) = 4.24$, $p < .001$ performed above chance but 3rd grade (M =2.44, SD =1.31) did not.

Inconclusive trials: 3rd grade: (M =1.60, SD =0.51; $t(15) = 4.39$, $p < .001$ and 5th grade: (M =1.40, SD =0.78; $t(17) = 2.12$, $p = .04$ were above chance, but 1st graders (M = 0.80, SD = 0.83) were not significantly different from chance.

2-Variable condition: 3 x 2 x 2 (Age × Gender x Judgment) ANOVA yielded a significant effect of Age Group, $F(2,47) = 5.13$, $p = .01$, partial $\eta 2 = 0.179$.



3 x 2 (Age × Experiment) ANOVA demonstrated a significant age effect, $F(2, 51) = 5.39$, $p = .007$, partial $\eta 2 = .175$.



3rd- and 5th- grade students performed significantly above chance for both confounded and unconfounded problems, 1st grade children did not.

Unconfounded trials
- 3rd grade: $t(15) = 2.53$, $p = .02$; 5th grade: $t(17) = 4.65$, $p < .001$

Confounded trials:
- 3rd grade: $t(15) = 5.71$, $p < .001$; 5th grade: $t(17) = 4.50$, $p < .001$

Summary and Conclusions

- Children across age groups performed above chance for evidence evaluation, but performance did not differ from chance for experiment evaluation. Evidence evaluation may be more accessible as it provides an outcome for children to analyze. Experiment evaluation required children to consider prospective outcomes which appears to be more difficult for children.

- Older children performed above chance for inconclusive 1-Variable tests, but younger children did not perform differently from chance. Thus, recognition of ambiguous experiments and evidence improved with age.

- For the 2-Variable condition, 3rd and 5th grade children successfully evaluated both confounded and unconfounded experiments, but 1st grade children did not. Reasoning about multiple variables and confounded experiments may require advanced skills and be challenging for younger children.

References

Chen, Z., & Klahr, D. (1999). All other things being equal: Children's acquisition of the control of variables strategy. *Child Development*, 70, 1098–1120.

Piekny, J., & Maehler, C. (2013). Scientific reasoning in early and middle childhood: The development of domain-general evidence evaluation, experimentation, and hypothesis generation skills. *British Journal of Developmental Psychology*, 31(2), 153-179.

Piekny, J., Grube, D., & Maehler, C. (2014). The development of experimentation and evidence evaluation skills at preschool age. *International Journal of Science Education*, 36(2), 334-354.

Sodian, B., Zaitchik, D., & Carey, S. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Child Development,* 62, 753–766.

Correspondence: tvilma1@niu.edu