

Children's inquiry learning strategies may depend on their prior conceptual knowledge

Lucas Lörch

lloerch@gmx.de

DIPF Leibniz Institute for Research and Information in Education <https://orcid.org/0000-0002-5665-3313>

Garvin Brod

DIPF Leibniz Institute for Research and Information in Education

Research Article

Keywords: scientific reasoning, science learning, prior knowledge, probabilistic learner modeling, Bayesian computational modeling

Posted Date: July 28th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10422716/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Abstract

The present study explored how children's prior conceptual knowledge affects their inquiry learning strategies. 139 elementary school children performed a pretest on water displacement. Then, they evaluated whether seeing the outcome of four different experimental setups would help them learning something about water displacement. The experimental setups compared two spheres of different sizes and materials and were designed in such a way that children's preferences indicated either an effect-producing or a controlled testing strategy. We conducted probabilistic modeling of children's prior beliefs and formed six belief clusters based on a hierarchical clustering algorithm. While children tended to evaluate all experiments as helpful, the pattern of preference indicated an effect-producing strategy for children with a clear misconception and a controlled testing strategy for children with an uncertain or correct belief. This finding, although exploratory, suggests that prior knowledge may explain why children sometimes succeed and sometimes fail in conducting informative experiments.

Introduction

There is an ongoing scientific debate about children's inquiry learning skills (see Lapidow & Walker, 2021). One line of research indicates that children have difficulties with inquiry learning tasks. That is, children were found to fail conducting informative experiments (Tschirgi, 1980; Zimmerman & Klahr, 2018), to seek generating confirmatory evidence (Croker & Buchanan, 2011; Klahr et al., 1993), and to lack a generalized understanding of the control-of-variables strategy (Zimmerman & Croker, 2013). In contrast, another line of research finds children to be like "intuitive scientists". That is, children were found being able to identify hypothesis tests that generate informative evidence (Bullock & Ziegler, 1999) and to identify confounded evidence in knowledge-lean contexts (Köksal et al., 2021). The partial incompatibility of these findings constitutes a "gap" in the debate about children's inquiry learning skills (see Kuhn & Pearsall, 2000; Schulz, 2012).

One potential explanation for this gap is the effect of prior knowledge. Prior knowledge can support learners in constructing hypotheses and testing them in a controlled manner (Schulz, 2012). However, prior knowledge can also irrationally bias learners towards those explanations they already deem plausible (Zimmerman & Klahr, 2018). As a consequence, learners may prefer experiments that have the potential to demonstrate the causal effect they expect, even if those experiments do not enable a controlled test of this effect. This approach has been labeled engineering desired outcomes (Schauble et al., 1991), effect-producing (Kuhn & Phelps, 1982), or positive testing (Köksal-Tuncer & Sodian, 2018). It is plausible that effect-producing depends on prior knowledge, as it refers to an effect that is expected by the learner beforehand. Despite its intuitiveness, the role of prior knowledge in children's inquiry learning has often been neglected.

Many previous studies avoided measuring prior knowledge but instead controlled the amount of prior knowledge by using knowledge-lean tasks, such as the blicket detector task (e.g., Köksal et al., 2021; Köksal-Tuncer & Sodian, 2018). Other studies used knowledge-rich tasks but operationalized prior

knowledge with sum-score measures and did not use these measures to explain individual differences (e.g., Croker & Buchanan, 2011; Tschirgi, 1980). Therefore, it is currently unclear whether and how inquiry learning strategies depend on prior knowledge.

To address these shortcomings, the present study used probabilistic learner models of children's prior knowledge. Based on the rational constructivist framework of cognition (Gopnik & Wellman, 2012; Xu, 2019), we conceptualized prior knowledge about science concepts as abstract, generative, and probabilistic causal beliefs. That is, we conceptualized prior conceptual knowledge as beliefs about general, inherently unobservable causal structures and as enabling predictions about the effects of these causal structures. We assumed that newly acquired beliefs do not replace intuitive beliefs, but that they exist in parallel (Potvin et al., 2015; Shtulman & Valcarcel, 2012). Based on these assumptions, we modeled prior conceptual knowledge as a categorical probability distribution that indicates the strength of each causal belief for each individual learner (Colantonio et al., 2023; Lörch et al., 2024).

We chose water displacement as the domain of conceptual learning. The advantage of this domain is that it is a knowledge-rich domain with established intuitive beliefs; elementary students typically believe mass or material rather than size to be the relevant causal variable (Linn & Eylon, 2000; Piaget & Inhelder, 1978). Moreover, learning experiments in this domain have already been developed (Brod et al., 2025; Theobald & Brod, 2021), and tested for informativeness (Lörch et al., 2024). In these learning experiments, children predict which of two spheres of varying sizes and materials displace more water. Then, they are shown an animation of the outcome which allows them to draw inferences about water displacement.

These experiments have been categorized into four types (see Table 1). In additive and contrastive comparisons, both the size and the material of the spheres vary in such a way that the larger sphere is made of the heavier or lighter material, respectively. In vary-only-size and vary-only-material comparisons, only the respective features are varied. Importantly, however, mass also varies between the spheres in these comparisons due to the inherent confoundedness of size, material, and mass (Lörch et al., 2024). An experiment can be considered informative if it disproves the mass and material beliefs and at the same time suggests the size belief. Thus, the mass and material beliefs and the size belief need to suggest different outcomes for an experiment to be informative. This is the case for contrastive and vary-only-material comparisons. In additive and vary-only-size comparisons, however, the mass and material beliefs and the size belief suggest the same outcome. Thus, these comparisons are uninformative (Lörch et al., 2024).

In our study, we employed an inquiry learning task. Children saw a sequence of experimental setups and evaluated whether seeing the outcome of these experiments would help them to learn something about water displacement. We chose the experimental setups in such a way that children's pattern of preference reflects their inquiry strategy. The experiments and the indicated inquiry strategy are shown in Table 1. Effect-producing in this context means that a learner wants to demonstrate the effect of mass or material. This would be reflected by a preference for the experiments in which the spheres clearly

differ in mass and material, i.e., the additive, contrastive, and vary-only-material comparisons. Applying the control-of-variables strategy (Chen & Klahr, 1999) in this context means that a learner wants to vary only one surface feature (size or material) at a time. This would be reflected by a preference for the vary-only-size and vary-only-material comparisons. Applying a controlled testing strategy in this context means that a learner wants to test the competing beliefs (size, mass, material) against each other. This would be reflected by a preference for experiments in which the size and the mass and material beliefs suggest different outcomes, i.e., the contrastive and vary-only-material comparisons.

We explored children's pattern of preference for the different experiments in relation to their prior beliefs about water displacement modeled from a pretest. In addition, we asked children why they expected the experiments to be informative. We expected that this mixed-method approach together with the probabilistic modeling of prior knowledge would provide novel insights on the role of prior knowledge in children's inquiry learning.

Method

The materials, data and analyses of the present study can be found on the Open Science Framework: <https://osf.io/j8c3g>

Sample

139 children aged 6-to-9 participated in the study ($M_{\text{age}} = 7.22$, $SD = 1.13$). 52% of the participants were male and 48% were female.

Procedure and Material

The experiment was conducted in individual sessions at a natural history museum in southwestern Germany. After parents provided informed consent, children were instructed that they would learn something about water displacement. They were shown a video in which a rubber ball was submerged, and they submerged a plastic cube in a glass of water themselves. This ensured that they understood that the relevant phenomenon was water displacement and not buoyancy. Moreover, they were instructed about the different objects used in the study: small, medium, and large spheres made from polystyrene, wood, or iron. They were handed equally sized spheres made from the three materials so that they could feel the difference in weight. Then, children performed a warm-up trial in which they predicted which of two spheres would displace more water and then saw an animation of the spheres being submerged.

After instruction and warm-up, children performed a 6-trial pretest. They made predictions which of two spheres would displace more water without seeing the outcome. Predictions were made on a three-point scale (1-left sphere displaces more water, 2-spheres displace equal amounts of water, 3-right sphere displaces more water; see Fig. 1).

Subsequently, children completed the experiment evaluation task. They saw the four experimental comparisons shown in Table 1 in random order and were asked whether the comparison would help them to learn something about water displacement. After choosing “yes,” “no,” or “don’t know,” they were asked why they considered this comparison helpful/not helpful (see Fig. 1). Their verbal response was noted by the experimenter.

Then, children saw a comparison randomly selected from the previous four, made a prediction and observed an animation of the outcome. They were asked whether this comparison did help them to learn something about water displacement (yes/no/don’t know) and what they had learned. Finally, we asked children directly about their current belief. Children saw an image of a large iron sphere displacing more water than a medium polystyrene sphere. We asked them why they thought that the large iron sphere displaced more water. All tasks were completed on a computer and the experimenter provided individual support in moving the mouse and providing responses.

Afterwards, children completed the German version of the Developmental Cognitive Reflection Test (CRT-D; Holstein et al., 2025). Note that the focus of this paper is the pretest and the experiment evaluation task. At the end of the experiment, children were debriefed and received a thank-you gift worth 5 €.

Analysis

Our analyses comprised 4 steps: (1) modeling children’s prior belief distributions, (2) performing hierarchical clustering on the prior belief distributions to determine prior belief clusters, (3) computing the proportion of children who evaluated the different comparisons as helpful in the different prior belief clusters and (4) performing thematic analysis with the open responses to determine which themes were dominant in each prior belief cluster.

To model children’s prior beliefs, we used the PrAM Python package[1] which is based on code presented in Colantonio et al. (2023). The modeling is based on Bayes’ rule, i.e., on the idea that the probability of a causal explanation after seeing some data (called the posterior) is proportional to the probability of that explanation before seeing the data (called the prior) multiplied with the probability of observing the data given the explanation were true (called the likelihood). In the PrAM package, this logic is translated to conceptual learning. The prior represents a learners initial causal belief, the posterior represents a learners belief after learning, and the likelihood represent the probability of observing some evidence given the belief were true.

We distinguished four beliefs about water displacement. According to the size, material, and mass beliefs, the respective variable determines water displacement. According to the random belief, water displacement is unsystematic. As explained above, we assumed that these beliefs exist in parallel in a learner’s mind. Thus, the belief distribution is defined as a categorical probability distribution that indicates the strength of each of these four beliefs for an individual learner (see bottom plots of Fig. 2 for an example).

We modeled children's prior beliefs about water displacement based on the outcomes they predicted in the pretest (for details, see Colantonio et al., 2023). The general logic is that each predicted outcome has a certain likelihood given each belief. For example, predicting that a small but heavy sphere displaces more water than a large but light sphere has a high likelihood given the mass belief, but a small likelihood given the size belief. Making this prediction thus increases the probability of the mass belief but decreases the probability of the size belief in this learner's prior belief distribution. Thus, to model children's prior beliefs, the package first computes the likelihoods of the different outcomes (i.e., left sphere displaces more water, spheres displace equal amounts of water, right sphere displaces more water) given each belief.

To enable this step, we coded the amount of displaced water in each pretest item given each belief with simple integers. For example, in the trial depicted in Fig. 1 (large wooden vs. medium iron sphere), the size belief suggests that the left sphere displaces more water. As there were three sizes (small, medium, large), and as the spheres in this trial were large and medium sized, we coded the amount of water displaced by the two spheres given the size belief as 3 and 2, respectively. For each pretest trial and each belief, the package then defines two normal distributions with the coded amount of displaced water as the mean. The package draws 1,000 values from these normal distributions and counts how often the value drawn from the first distribution is larger, the

value drawn from the second distribution is larger, or the values are similar. Translating these counts into proportions provides the likelihoods of outcomes given each belief.

Then, the package determines the likelihood of each child's responses in each pretest trial given each belief and multiplies the likelihoods of the different trials. Thereby, the package computes the joint probability of each child's sequence of responses given each belief, which forms the prior belief distribution.

For our hierarchical clustering algorithm, we computed the Jensen-Shannon Divergence (JSD) to indicate the dissimilarity between all combinations of prior belief distributions. Then, we used Ward's method to perform hierarchical clustering (Murtagh & Legendre, 2014; Žibera et al., 2004). This method, in simplified terms, creates a hierarchy of clusters by minimizing the within-cluster variance in a stepwise manner. We explored the interpretability of the prior belief distributions for different ways of clustering. We sought for a way of clustering yielding a minimal number of clearly interpretable clusters.

Results

Figure 2 shows the results of our exploratory analysis of the hierarchical clustering of prior beliefs. We chose a solution with six clusters: Four clusters contained children with a single strong belief ("1-materialBelief", "3-sizeBelief", "4-massBelief", "5-randomBelief"), two clusters contained children who were undecided between multiple beliefs. One cluster contained children who were undecided between the mass and material beliefs ("2-undecidedMassMaterial"), while the other contained children who were undecided between the size belief and the mass and random beliefs ("5-undecidedSizeMass"). We

decided not to present the results for the random belief cluster (cluster 5), as the 15 participants in this cluster mainly responded unsystematically in the pretest and this could indicate that these children were distracted or had troubles understanding the task.

Figure 3 shows the mean percentage of children who evaluated a comparison as helpful, separately for the four comparisons and five belief clusters. We see that the mean percentage never drops below 65%, that is, children generally tended to evaluate the experiments as helpful. However, we also see that there are clear differences in children's preferences between the four experiments. The top plot shows that the pattern of preference for the size and material belief clusters indicates an effect-producing approach. Participants with a clear mass or material belief preferred those comparisons in which there was a clear difference in mass or material between the spheres. This pattern was more pronounced in the material belief cluster than in the mass belief cluster. The notion of effect-producing was also a recurring theme in children's reasoning why they considered certain comparisons to be helpful. In their reasoning, children frequently stated that a comparison was useful because there was a difference in mass or material between the spheres. For example, Child 114 stated that the vary-only-material comparison was useful „because the right ball is heavier.“ Moreover, rejecting the vary-only-size comparison was often justified with a lacking difference in mass or material between the spheres. For example, child 36 argued that the vary-only-size comparison was not informative because „both [spheres] weight the same.“ Likewise, Child 104 stated that the vary-only-size comparison is not helpful because “both [spheres] have the same material, so they have the same weight and you cannot understand that.” We see that both the quantitative and the qualitative data suggests that children with a clear misconception chose an effect-producing strategy.

For the size belief cluster, the top plot of Fig. 3 indicates a controlled testing approach. Participants with a size belief preferred those comparisons that allowed testing the size belief against the mass and material beliefs. We also found some support for a controlled testing approach in the open responses. Some children explained in a surprisingly sophisticated way that the additive comparison was not helpful because it confounded size and mass. For example, Child 132 claimed, “because [the] iron [sphere] is much larger and heavier than the polystyrene sphere, the [additive comparison] does not make any sense”. However, the qualitative evidence was less clear for the size belief cluster than for the mass and material belief clusters.

Lastly, for the undecided belief cluster, the bottom plot of Fig. 3 also indicates a controlled testing approach. However, the pattern is less clear than in the size belief cluster. More specifically, participants in the undecided belief clusters did not reject the vary-only-size comparison as clearly as those in the size belief cluster. This suggests that

participants with an undecided belief might have been less aware of the “hidden” confoundedness of size and mass in the vary-only-size comparison. The clear rejection of the additive comparison and the preference for the highly informative contrastive comparison nevertheless support the notion of a controlled testing approach. Some children in the undecided belief clusters also provided sophisticated

reasons for the confoundedness of the additive comparison and for the informativeness of the contrastive comparison. Such argumentation is best reflected by Child 96 who stated that the additive comparison is not informative because „small and light versus large and heavy does not help,“ and that the contrastive comparison is informative „because I [can] learn if mass or size displaces more water“. In summary, the quantitative data suggests that children in the size and undecided belief clusters chose a controlled testing approach and there is some support for this conclusion in the qualitative data.

Discussion

There is an ongoing scientific debate about the inconsistencies in previous findings on children's inquiry learning skills (see Lapidow & Walker, 2021). We believe that the understanding of these inconsistencies can be advanced by clarifying the role of prior knowledge in inquiry learning. We chose a novel methodological approach that combined probabilistic modeling and hierarchical clustering of children's prior beliefs with qualitative data on children's reasoning.

Our study provides first evidence that children's inquiry approaches may depend on their prior conceptual knowledge. If children have a strong intuitive belief, they may aim to demonstrate the correctness of this belief by producing an expected effect (cf. Croker & Buchanan, 2011; Tschirgi, 1980). If children are uncertain in their beliefs, they may rather test hypotheses in a controlled manner (cf. Bullock & Ziegler, 1999). Due to their uncertainty, these children may predict the outcome for each of their competing beliefs. This may help them to recognize if the outcomes for competing beliefs are the same in some comparison and that it thus does not yield any helpful insights.

Interestingly, we found the clearest evidence for a controlled testing approach for children with a correct size belief. These children apparently did not intend to produce size effects. So merely concluding that a strong belief promotes effect-producing would be an unwarranted simplification. Which differences between the size belief and the mass and material beliefs may have caused these differences in behavior? It can reasonably be assumed that the mass and material beliefs are well-established, intuitive beliefs that most children acquired through their everyday experiences and held at some point in their development (Piaget & Inhelder, 1978). The size belief, however, may have been less well-established. It is likely that children had to explicitly learn this belief by overcoming their initial mass or material belief. This would mean that they were aware that they *used to believe* that mass or material determines water displacement but that they *now believe* that it is indeed size. Based on this line of argumentation, it seems plausible that children with a size belief in our study wanted to reassure their recently acquired belief and thus preferred informative experiments.

This finding raises numerous questions that should be addressed by future studies. Would children revert to producing size effects if the size belief was more well-established? Or is there a fundamental epistemic difference between intuitive and consciously learned beliefs? And is there perhaps a meta-level certainty in one's beliefs that is not yet captured by our probabilistic modeling? Investigating these questions would deepen our understanding of conceptual knowledge and its role in inquiry learning.

In none of the belief clusters did children prefer comparisons in which only one surface feature (size or material) varied between spheres. That is, children were not focused on applying the control-of-variables strategy, in line with previous findings (Zimmerman & Croker, 2013). Nevertheless, children were sensitive to confoundedness (cf. Köksal et al., 2021), at least when it occurred between surface features of the spheres. That is, children frequently recognized the confoundedness of size and material in additive comparisons, but rather struggled to recognize the “hidden” confoundedness of size and mass in vary-only-size comparisons. This suggests that learning contexts that entail inherent confounds such as the one between size and mass may be challenging for children.

In addition to these theoretical implications, our study revealed the potential of probabilistic learner models as a method for understanding prior knowledge effects. Conceptual knowledge has been conceptualized as abstract, generative, and probabilistic (Gopnik & Wellman, 2012; Xu, 2019). Probabilistic learner models account for these characteristics and therefore may reveal effects that other methods miss. Moreover, clustering algorithms enable performing group-level analyses with the individual-level data that result from the probabilistic learner modeling. It should be noted, however, that the results of the probabilistic learner models are dependent on the specific items used in the pretest. Experimental comparisons differ in their potential to distinguish certain beliefs. Thus, the experimental comparisons contained in a pretest affect the resulting prior belief distributions and the meaning of the clusters. Probabilistic learner models do not reveal some objective truth but should be seen as a tool providing a new perspective on learners.

As a limitation to the results of our study, it should be noted that the differences in the preferences for the four experiments were rather small. Children generally tended to evaluate all experiments as helpful. This may have also been caused by our task design, as there were no costs of evaluating experiments as helpful. Future studies could present two experimental setups to children and ask which outcome they would like to observe. This would provide a clearer measure of preference. Moreover, children sometimes struggled with the hypothetical nature of our questions, and the open responses were often hard to interpret. Lastly, our study was exploratory and the pattern of effects was rather complex. The main finding was an interaction between five belief clusters and four types of experiments. Future studies should perform a pre-registered experimental test of our assumptions and should either aim for sample sizes sufficient to reveal complex interactions, or test less complex effect patterns.

In summary, combining probabilistic learner models and qualitative data, our study found that children’s inquiry strategies depended on their prior conceptual knowledge. Children with a single, incorrect intuitive belief tended to demonstrate the correctness of this belief. Children with uncertain beliefs were able to identify experiments that could help them to gain more certainty. Prior knowledge could thus explain why children sometimes succeed and sometimes fail when conducting informative experiments.

Declarations

the study was approved by the ethics committee of the DIPF | Leibniz Institute for Research and Information in Education.

References

1. Brod G, Holstein E, Weindorf L, Colantonio J, Bonawitz E, Theobald M (2025) Do it yourself: Discerning the effects of self-directed activity on conceptual learning. *Npj Sci Learn* 10(1):70. <https://doi.org/10.1038/s41539-025-00364-9>
2. Bullock M, Ziegler A (1999) Scientific reasoning: Developmental and individual differences. In F. E. Weinert & W. Schneider (Hrsg.), *Individual Development from 3 to 12: Findings from the Munich Longitudinal Study* (S. 38–54). Cambridge University Press. <https://www.semanticscholar.org/paper/Scientific-reasoning%3A-Developmental-and-individual-Bullock-Ziegler/2838ce890d0fc4f40097562f4af7a15b946f09fa>
3. Chen Z, Klahr D (1999) All Other Things Being Equal: Acquisition and Transfer of the Control of Variables Strategy. *Child Dev* 70(5):1098–1120. <https://doi.org/10.1111/1467-8624.00081>
4. Colantonio J, Bascandziev I, Theobald M, Brod G, Bonawitz E (2023) Priors, Progressions, and Predictions in Science Learning: Theory-Based Bayesian Models of Children’s Revising Beliefs of Water Displacement. *IEEE Trans Cogn Dev Syst* 15(3):1487–1500. <https://doi.org/10.1109/TCDS.2022.3220963>
5. Croker S, Buchanan H (2011) Scientific reasoning in a real-world context: The effect of prior belief and outcome on children’s hypothesis-testing strategies. *Br J Dev Psychol* 29(3):409–424. <https://doi.org/10.1348/026151010X496906>
6. Gopnik A, Wellman HM (2012) Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychol Bull* 138(6):1085–1108. <https://doi.org/10.1037/a0028044>
7. Holstein ER, Lörch L, Brod G (2025) Children’s Cognitive Reflection Skills Predict Their Conceptual Learning. *Collabra: Psychol* 11(1):143423. <https://doi.org/10.1525/collabra.143423>
8. Klahr D, Fay AL, Dunbar K (1993) Heuristics for Scientific Experimentation: A Developmental Study. *Cogn Psychol* 25(1):111–146. <https://doi.org/10.1006/cogp.1993.1003>
9. Köksal Ö, Sodian B, Legare CH (2021) Young children’s metacognitive awareness of confounded evidence. *J Exp Child Psychol* 205:105080. <https://doi.org/10.1016/j.jecp.2020.105080>
10. Köksal-Tuncer Ö, Sodian B (2018) The development of scientific reasoning: Hypothesis testing and argumentation from evidence in young children. *Cogn Dev* 48:135–145. <https://doi.org/10.1016/j.cogdev.2018.06.011>
11. Kuhn D, Pearsall S (2000) Developmental Origins of Scientific Thinking. *J Cognition Dev* 1(1):113–129. https://doi.org/10.1207/S15327647JCD0101N_11
12. Kuhn D, Phelps E (1982) The Development of Problem-Solving Strategies. In H. W. Reese (Hrsg.), *Advances in Child Development and Behavior* (Bd. 17, S. 1–44). JAI. <https://doi.org/10.1016/S0065->

13. Lapidow E, Walker CM (2021) Rethinking the gap: Self-directed learning in cognitive development and scientific reasoning. *Wires Cogn Sci* 13(2). <https://doi.org/10.1002/wcs.1580>
14. Linn MC, Eylon B-S (2000) Knowledge Integration and Displaced Volume. *J Sci Edu Technol* 9(4):287–310. <https://doi.org/10.1023/A:1009451808539>
15. Lörch L, Bonawitz E, Brod G (2024) Challenging the control-of-variables strategy: How confounded comparisons can support children’s science learning. In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Hrsg.), *Proceedings of the Annual Meeting of the Cognitive Science Society* (Bd. 46, S. 1054–1060). <https://escholarship.org/uc/item/42c70920>
16. Murtagh F, Legendre P (2014) Ward’s Hierarchical Clustering Method: Clustering Criterion and Agglomerative Algorithm. *J Classif* 31(3):274–295. <https://doi.org/10.1007/s00357-014-9161-z>
17. Piaget J, Inhelder B (1978) *Le développement des quantités physiques chez l’enfant: Conservation et atomisme* (4 éd). Delachaux et Niestlé
18. Potvin P, Sauriol É, Riopel M (2015) Experimental evidence of the superiority of the prevalence model of conceptual change over the classical models and repetition. *J Res Sci Teach* 52(8):1082–1108. <https://doi.org/10.1002/tea.21235>
19. Schauble L, Klopfer LE, Raghavan K (1991) Students’ transition from an engineering model to a science model of experimentation. *J Res Sci Teach* 28(9):859–882. <https://doi.org/10.1002/tea.3660280910>
20. Schulz L (2012) The origins of inquiry: Inductive inference and exploration in early childhood. *Trends Cogn Sci* 16(7):382–389. <https://doi.org/10.1016/j.tics.2012.06.004>
21. Shtulman A, Valcarcel J (2012) Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition* 124(2):209–215. <https://doi.org/10.1016/j.cognition.2012.04.005>
22. Theobald M, Brod G (2021) Tackling Scientific Misconceptions: The Element of Surprise. *Child Dev* 92(5):2128–2141. <https://doi.org/10.1111/cdev.13582>
23. Tschirgi JE (1980) Sensible Reasoning: A Hypothesis about Hypotheses. *Child Dev* 51(1):1–10. <https://doi.org/10.2307/1129583>
24. Xu F (2019) Towards a rational constructivist theory of cognitive development. *Psychol Rev* 126(6):841–864. <https://doi.org/10.1037/rev0000153>
25. Žiberna A, Kejžar N, Golob P (2004) Comparison of different approaches to hierarchical clustering of ordinal data. *Adv Methodol Stat* 1(1):57–73. <https://doi.org/10.51936/tivd4881>
26. Zimmerman C, Croker S (2013) Learning science through inquiry
27. Zimmerman C, Klahr D (2018) Development of Scientific Thinking. In J. T. Wixted (Hrsg.), *Stevens’ Handbook of Experimental Psychology and Cognitive Neuroscience* (1. ed., pp. 1–25). Wiley. <https://doi.org/10.1002/9781119170174.epcn407>

Footnotes

Table 1

Table 1 is available in the Supplementary Files section.

Figures

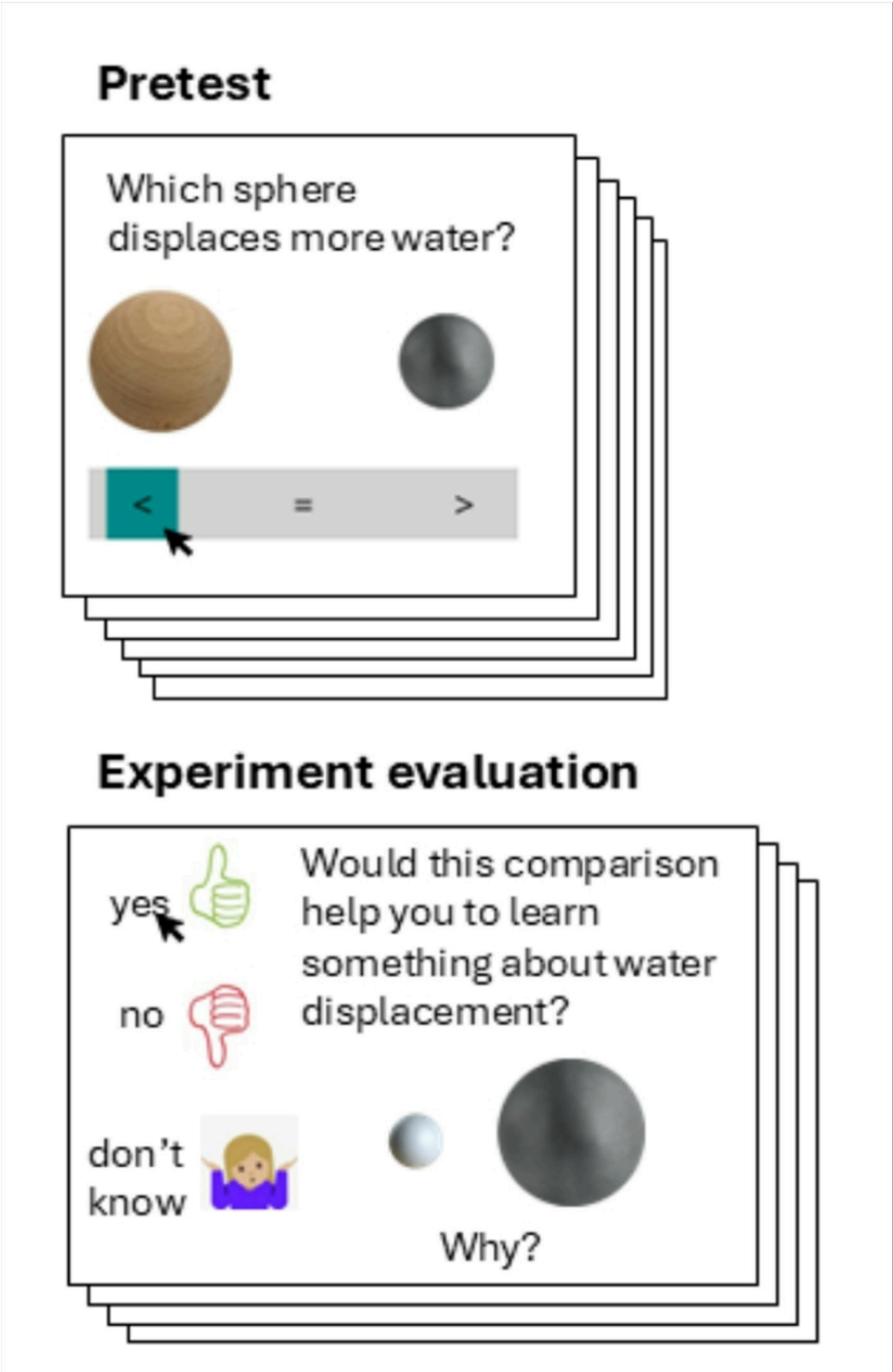


Figure 1

Pretest and experiment evaluation task

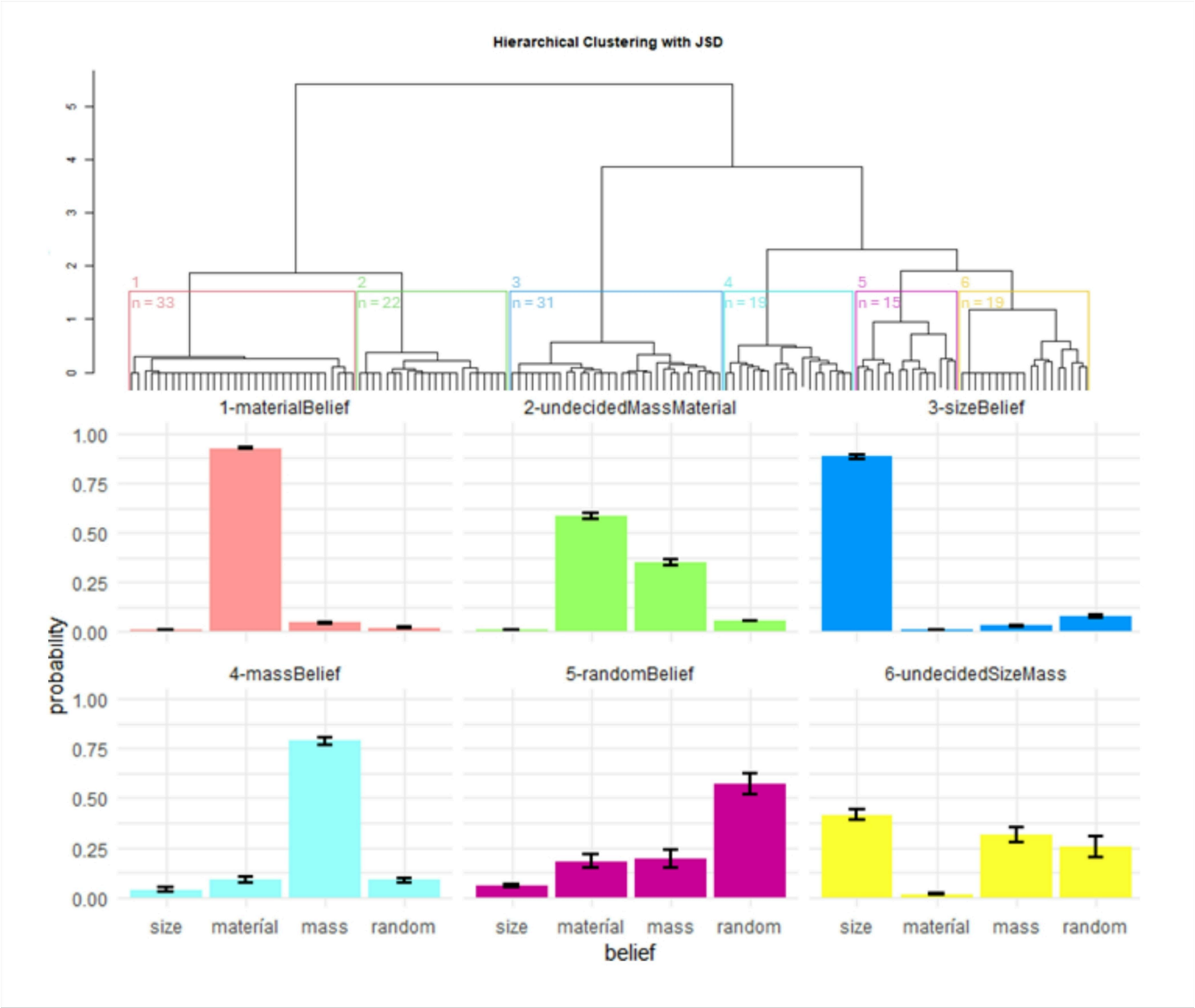


Figure 2

Clustering of prior beliefs in the present study. The top plot shows the dendrogram with the hierarchical clustering structure. The leaves on the bottom of the tree represent the prior belief distributions of the individual participants. The tree represents the hierarchical clustering based on the dissimilarity of these distributions. We decided to divide our sample into six clusters indicated as colored rectangles. The bar plots on the bottom show the mean prior belief distributions and labels of each cluster.

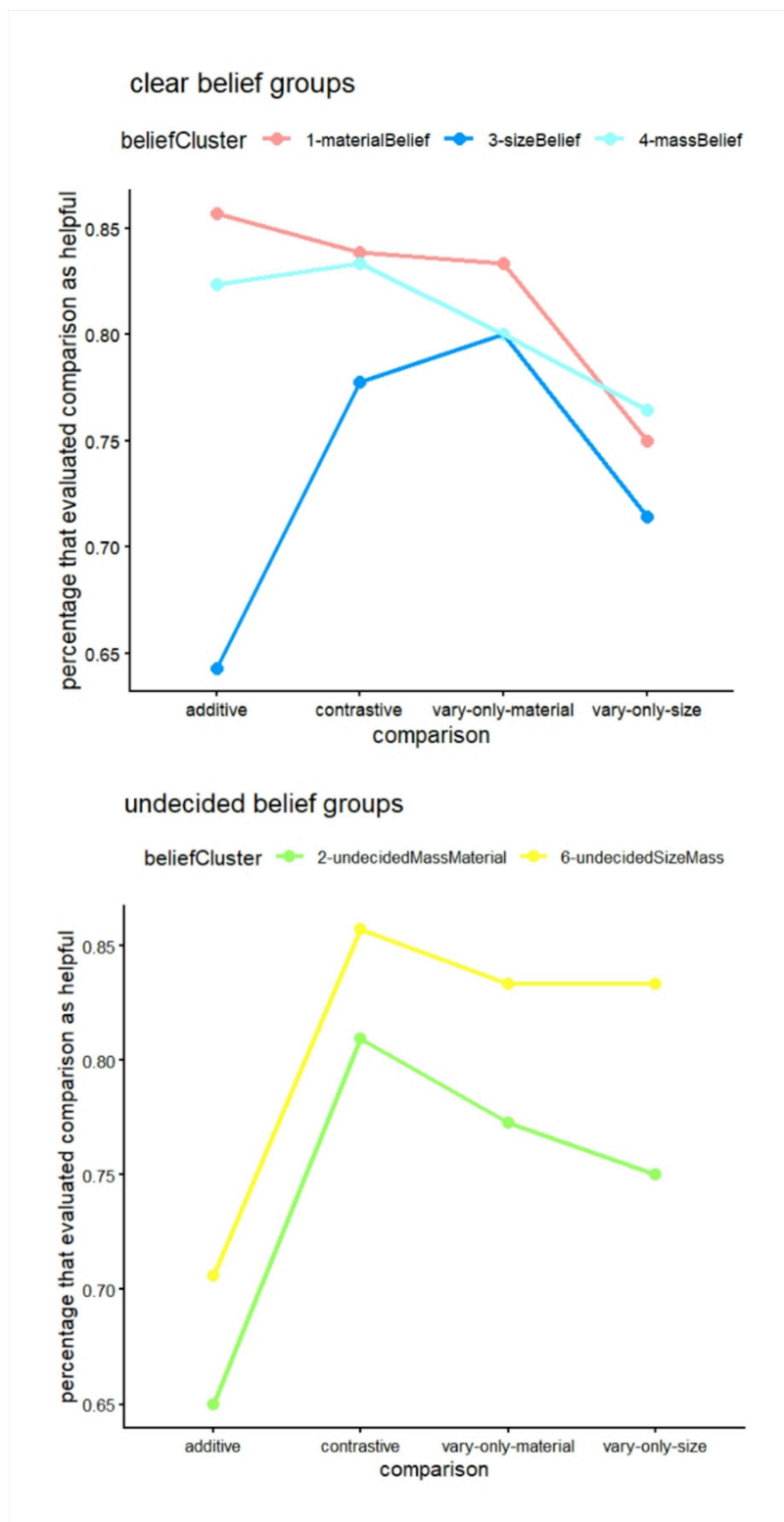


Figure 3

Mean percentage of children that judged a comparison as helpful, separately for the four comparisons and five belief clusters.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table1.docx](#)