Preschoolers' Learning about Buoyancy: Does it help to give away the answer?

Heather Baker[1], Anna Haußmann[1], Heidi Kloos[1], & Anna Fisher[2]
[1]Department of Psychology, University of Cincinnati, Cincinnati OH, 45221
[2]Department of Psychology, Carnegie Mellon University, Pittsburgh PA, 15213

Abstract—**How to teach science to young children? Despite extensive research on the importance of direct instruction, the general recommendation is to provide children with an opportunity to explore a domain on their own. However, it is not clear how self-guided explorations could help children overcome misconceptions. The current paper pursues this question within the domain of sinking objects – a domain in which children often hold misconceptions. An open-exploration teaching approach was contrasted with a teaching approach in which children were provided with a simplified rule about buoyancy. In particular, preschoolers were taught to focus on the amount of empty space inside a sinking object: the less empty space the faster it would sink. Results of two experiments show some benefit of the direct-instruction approach. Although performance was not at ceiling, giving away the answer nevertheless improved performance. It might have allowed children to circumvent their misconceptions, setting the stage for further learning. Implications for science of learning in humans and robotics are discussed.**

Index Terms — **Buoyancy, Direct Instruction, Science Learning, Self-guided exploration.**

I. INTRODUCTION

In science education direct instruction has a bad reputation, especially when it comes to young children (e.g., [1], [2]). In fact, direct instruction is commonly seen as the very culprit of often-found difficulties in science learning. Whether shortcomings in science learning are measured as standardized test scores and compared across nations, or in terms of a child's ability to transfer knowledge to a new domain, direct instruction is rarely embraced as a way out. Instead, the promising alternative in science education is seen in self-guided explorations, hands-on activities, and inquiry, the activity of asking questions and answering them through experimentation (e.g., [3], [4], [5]).

The bad reputation of direct instruction is surprising, given a large body of empirical evidence in support of direct instruction in both human (e.g., [6], [7], see also [8]) and robotic (e.g. [9]) learning. Furthermore, even though exploratory science learning approaches are viewed as having educational potential (e.g. [10]) and are unquestionably attractive and easily accepted by the student, there are theoretical reasons to believe that self-guided explorations, hands-on activities, or inquiry could hurt successful science learning. This is because scientifically valid concepts are often far less salient than irrelevant surface features. For example, when asked to predict objects' rate of sinking, the correct feature of density is much less salient than the feature of heaviness. For learning to happen, one needs to attend to such hidden features and ignore irrelevant – but nevertheless salient – features. When left without guided support, this is quite difficult even for adults. Indeed, even though it has been argued that given enough time during self-guided explorations, a child might eventually discover the rules behind a physical phenomenon like density on their own, when exploring sinking objects without guidance children up to 12 years of age focus on variation of mass or volume alone, failing to detect the less salient relation of density (e.g., [11], [12], [13]). For example, children demonstrate the mistaken belief that heaviness alone predicts the sinking behavior of objects. Most likely, self-guided explorations then lead to confirming such mistaken beliefs, rather than challenging them (cf. [14]).

The goal of the current study is to further explore the effect of direct instruction in science education. As the strongest case, we look at science education at preschool level, a setting in which science is often explicitly limited to explorations only (e.g., [15]). Could preschoolers benefit from direct instruction to understand a science concept? To address this question, we focused on the science of how density affects the sinking behavior of objects.

A. Density and Sinking Objects

Density has visible effects on the behavior of objects: denser objects sink faster than less dense objects; and objects with a density smaller than that of water float. This makes density an ideal science topic to be explored in a self-guided manner: Even young children can experience sinking and floating objects on their own, say at a water table in their classroom. Indeed, while density can be rather complex, its effect on sinking behavior can be simplified as the relative amount of empty space inside the object: The less empty space, the faster the object will sink – a relation that even preschoolers can discover spontaneously ([16]).

Furthermore, even though density is a common topic in middle school and high-school science curricula, misconceptions about sinking objects often remain unchanged even after formal instructions (e.g., [12]). For example, despite controlled demonstrations about the difference between mass and density, children nevertheless hold onto their mistaken idea that the mass of objects maps perfectly onto their sinking behavior ([17]). Given that these misconceptions are stable across age, it is pertinent to determine how they can be changed early on.

The current study builds upon these considerations and investigates the extent to which direct instruction could improve preschoolers' understanding of density. In two studies, preschoolers participated in one of two conditions that differed in the degree to which children were provided with top-down information about what feature to focus on. In the *give-away-the-answer* condition, the researcher taught children to focus on the amount of empty space inside the object. This training either was completed with the very objects used during assessment (Experiment 1), or it was completed with a novel set of objects, assessing possible transfer (Experiment 2). In the *open-exploration* condition (Experiment 1), children explored sinking objects on their own for a short period of time. In the *control* condition (Experiment 2), children participated merely in the assessment (identical across all conditions): Their task was in all cases to pick the faster sinking object from a pair of two objects. Importantly, objects within a pair differed in mass and volume in such a way that correct performance across pairs could not stem from focusing on mass or volume alone. The heavier of the two objects was faster on some trials, but not others. The larger of the two was faster on some trials, but not others. This choice of test items made it possible to determine the degree to which children base their answers on the wrong feature.

II. EXPERIMENT 1

A. Method

Participants. Participants included 20 3- to 5-year-old preschoolers of mixed racial background from local Head Start centers (mean age = 52 months; 7 boys and 13 girls). Half of the children participated in the *give-away-the-answer* condition, and half of the children participated in the *open-explorations* condition.

Materials. Materials consisted of 18 clear-glass jars with black lids that differed in their size (see Figure 1). Ten jars were 8 cm high and 6.3 cm wide; four jars were 6.9 cm high and 5.8 cm wide, and three jars were 5 cm high and 5.3 cm wide. Round aluminum discs (1 cm high, 4 cm in diameter, 43 g) could be placed inside the jars to obtain a desired mass. Large jars could hold between one and five discs (ranging in density between 0.81 and 1.54g/cm^3); medium jars could hold between one and four discs (ranging in density between 0.94 and 1.68 g/cm^3); and small jars could hold between one and three discs (ranging in density between 1.21 and 2.06 g/cm^3).

Three of the jars were used for training: a large one with one disc (190 g/cm^3), a medium jar with two discs (202 g/cm^3), and a small jar with two discs (170 g/cm^3). They could be dropped in a clear plastic container filled with water.

For testing, jars were combined into 24 pairs, such that one jar in a pair was always heavier than the other, one jar was always larger than the other, and one jar was always denser than the other. Figure 1 shows some example combinations. For eight pairs, the denser jar was heavier and larger (Figure 1A), for another eight pairs, the denser jar was heavier and smaller (Figure 1B), and for the remaining eight pairs, the denser jar was lighter and smaller (Figure 1C).

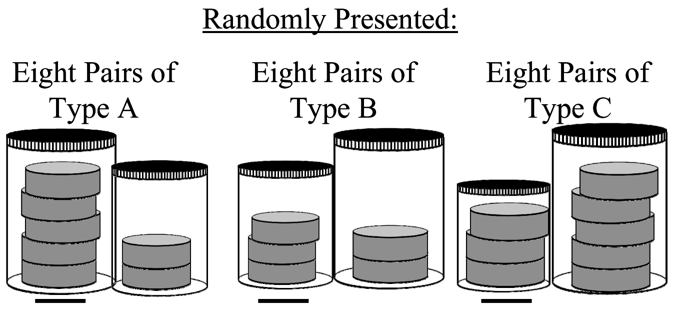

Figure 1. Schematic of three pairs of jars with weights. Relative to the losing jar, the faster sinking jar (marked with a vertical line underneath) could be heavy and large (A), heavy and small (B), or light and small (C).

Procedure. Children in this and the next experiment were tested by the experimenter in a one-on-one setting in a quiet area at their school. The experiment consisted of a preliminary 3-minute training phase (which differed as a function of condition), and a subsequent 10-minute assessment (which was identical for all children).

During training, children were introduced to the three training jars and the container filled with water. In particular, children in the *open-exploration* condition were shown all three jars, and they were told: "Here are some jars with weights inside of them, and a bucket of water. For a few minutes I am going to let you play with them." Children were encouraged to hold the jars, drop them into the water, and explore what makes one sink faster than the other. For example, the experimenter said: "Let's see what happens when you put the jar in the water. I wonder why that one stayed on top of the water and the other didn't? ". In contrast, children in the *give-away-the-answer* condition were presented with one pair of jars at a time and told: "Here are two jars with weights inside of them and a bucket of water. Which one do you think will sink faster?" The experimenter then dropped the pair of jars in the bucket explaining: "This jar sank fastest because there is less empty space inside it. Look at these two jars. See the empty space around the weights? This one has just a little tiny bit of empty space and this one has a lot of empty space. The jar that will sink faster is the one that has less empty space inside it."

During testing, participants were given one pair of testing jars at a time (with counterbalanced left-right placement and in random order). The pair was placed on the table in front of the participant, with the instruction to determine which jar would sink faster in water. Participants were encouraged to hold, feel, and look at the jars to make their best guess. No feedback was provided, and participants were not reminded of the rule about empty space.

B. Results and Discussion

To reflect children's success in the prediction task, three proportion-correct scores were calculated, one for each type of jar pair. Figure 2 shows the means of these scores as a function of condition and pair type (whether the winner in a pair was heavy and big, heavy and small, or light and small).

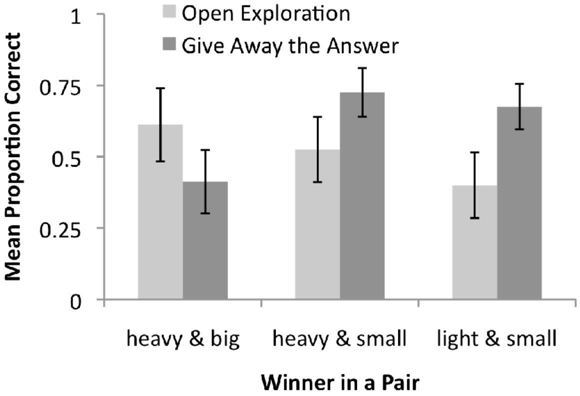

Figure 2. Mean proportion of correct choices during the testing phase, separated by condition and pair type. Standard errors are represented as error bars.

Between-group t-tests were conducted for each separate trial type. They revealed a reliable effect of condition when the winner was light and small, $t(18) = 3.1$, $p < .05$, in that children in the *give-away-the-answer* condition ($M = 0.68$) outperformed children in the *open-exploration* condition ($M = .40$). A marginally reliable effect of instruction was found when the winner was heavy and small, $t(18) = 2.1$, $p = .08$, suggesting that the brief prompts about the relevant feature of empty space were more beneficial than presenting children with the same objects in a self-exploration setting. Surprisingly, however, when the winner was heavy and big, children in the *give-away-the-answer* condition did not outperform children in the *open-exploration* condition ($t <$ 1,4). In fact, though not reliably different, mean proportion of correct choices was higher in the latter ($M = .61$) than the former condition ($M = .41$). Experiment 2 sought to test the robustness of the observed effects with a new group of preschoolers. We also included a group of adults, to determine ideal performance in this task. Finally, we included a control group in which participants only received the assessment, without any kind of prior training.

III. EXPERIMENT 2

Preschoolers and adults participated in one of two conditions: a *give-away-the-answer* condition and a *control* condition. Participants in the *control* condition were merely presented with the prediction task, allowing us to gain a better understanding of the task. Participants in *the give-away-the-answer* condition participated in a prior training phase that differed in one important way from the corresponding condition of Experiment 1: jars used during testing were different from the objects used during training. In fact, training was conducted with non-transparent cubes that either floated or sank in water. Participants were taught to imagine the 'insides' of the cubes such that they could evaluate the amount of empty space inside. This change in procedure allowed for the evaluation of both: children's ability to transfer their newly gained knowledge to a new context, and their ability to abstract the rule about empty space.

A. Method

Participants. Participants included 20 5-year-olds (10 girls, 10 boys; mean age = 64 months) recruited from urban and suburban middle-class day care centers and elementary schools. Adults ($N = 17$; mean age = 20.21 years; 3 women, 14 men) were included as a control group. They were recruited through a subject pool of introductory psychology classes and received partial course credit for participation.

Materials. The materials used for the assessment were the same as in Experiment 1: Pairs of jars differed in mass, volume, and density, such that the fastest sinking jar was heavier and larger in eight pairs, heavier and smaller in another eight pairs, and lighter and smaller in the remaining eight pairs.

Different from Experiment 1, training materials consisted of wooden cubes of various sizes, hollowed out and filled with lead to achieve a certain density. In particular, cubes either sank in water (density ≅ 2.0g/cm^3), or they floated (density ≅ 0.5 g/cm^3). There were 10 cubes in total (5 sinkers, 5 floaters), four of which were used during feedback training and six of which were used during non-feedback training. All cubes were painted the same color, such that no single cube could be differentiated from another. Table 1 provides information about mass and volume for each of these cubes.

Table 1. Dimensions of mass (in g) and volume (in cm^3) of the cubes used during training.

	High Density		Low Density	
	Mass	Volume	Mass	Volume
Feedback Training				
	78	45.6	200	113.5
	125	66.3	320	171.5
No-Feedback Training				
	31.8	62.5	81.5	171.5
	50.25	108	127.8	265. 8
	509	270.7	202	402.2

A so-called X-ray box was created, made out of a trapezoidal box (24.5 cm high; 17 cm wide on the top; 21.5 cm wide on the bottom). The right side of Figure 3 shows a schematic of the box. A blue light was placed at the bottom of the box controlled by a switch outside of the box and connected to an electrical cord. A horizontal see-through platform was mounted above the light such that cubes could be placed on top of it, and the light from below would illuminate the cube. Above the platform, the front and right side of the box were made out of clear plastic in order to provide visual access to the cube inside the box. The left side had a flap-like door for the cubes to be placed through. A USB cord was threaded through the box and was plugged into the computer to give the impression that box was actually connected to the laptop.

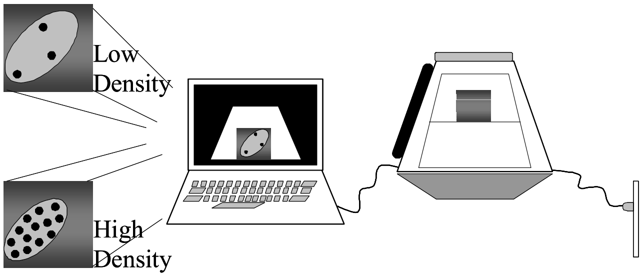

Figure 3. Schematic of setup used during the *give-away-the-answer* condition of Experiment 2.

To convey the insides of cubes, computer-generated images were used to show the ‘insides’ of the cube placed in the machine. The computer-generated images consisted of a square (to represent the outline of the cube) and an egg shaped ‘window’ with black dots (to represent the space inside the cube). Figure 3 shows two of such images, one representing a floater cube (with few dots), and one representing a sinker cube (with many dots).

Procedure. The cover story for both conditions involved a character named Suzie the Scientist, depicted on the computer screen.

In the *give-away-the-answer* condition, participants heard that Suzie the Scientist liked to know what sinks the fastest in water. Participants were told (while being shown corresponding images on the computer): “What matters is the empty space on the inside. If there is only a little bit of empty space, the cube will sink really fast. So Susie needs to look at how much empty space is on the inside of a cube. Suzie built a machine so we can see the inside of the cubes.” The experimenter then introduced the X-ray machine, turned on its light, and plugged in the USB cord into the side of the laptop. Next, participants were presented with one of the feedback-training cubes (randomly chosen from the set of four cubes). They were asked to hold and feel the cube, after which it was placed inside the X-ray machine. Once the flap was closed on the machine, the experimenter showed the cube’s ‘insides’ on the computer. For a high-density cube, the experimenter explained: “There are a lot of dots in this cube, so there is just a little bit of empty space around the dots. This means that it sinks faster.” For low-density cubes, the experimenter explained: “There’s a lot of empty space around the dots, so that means it sinks slower.” The cube was then removed from the machine, and participants were encouraged to hold and feel it again. While the participants were holding the cube, the instructions were: “Can you tell it has a lot of/just a little empty space?” The same steps were repeated with all the remaining feedback cubes.

The experimenter then removed the machine from sight, and four feedback trials followed (in randomized order). For each of the four trials, participants were presented with one of the cubes at a time, and they had to decide whether its insides matched with a picture of a high-density cube or a low-density cube (displayed on the computer). The instructions were: “Remember the cube that doesn’t have a lot of empty space around the dots will sink the fastest, and the one that has a lot of empty space around the dots will not sink the fastest, it might even float. Feel this cube. Do you think it looks like this or this on the inside?” Feedback was provided based off of the participant’s answer.

Finally, to assess whether participants could remember the cube training, six more cube trials followed, this time without feedback. Again, participants had to decide whether a cube had the insides of a cube with a lot of empty space, or whether it had the insides of a cube with very little empty space. They were reminded (1) that things with just a little bit of empty space will sink really fast in water, and (2) that things with a lot of empty space sink really slowly in water.

Jar assessment started immediately after the six cube trials without feedback. At this time, participants were told that Suzie the Scientist had some jars and needed to know which would sink the fastest in water. Participants were also instructed: “Just like the cubes, you can see the empty space in the jars. A jar with a lot of empty space will not sink very fast in water. But a jar that has only a little empty space will sink very fast in water”. The procedure of the ensuing assessment was identical to the assessment used in the *control group* and identical to the assessment used in Experiment 1.

The *control group* was simply presented with the identical jar assessment cover story as the *give away the answer* condition, which explained that Suzie the Scientist has some jars and needs to know which would sink the fastest in water. They were then given the jar assessment.

B. *Results and Discussion*

Figures 4 (adults) and 5 (preschoolers) show participants’ performance in each of the three pair types (whether the winner in a pair was heavy and big, heavy and small, or light and small), represented as mean proportion-correct scores, as a function of age group (adults vs. preschoolers) and type of instruction (control/no instruction vs. give-away-the-answer).

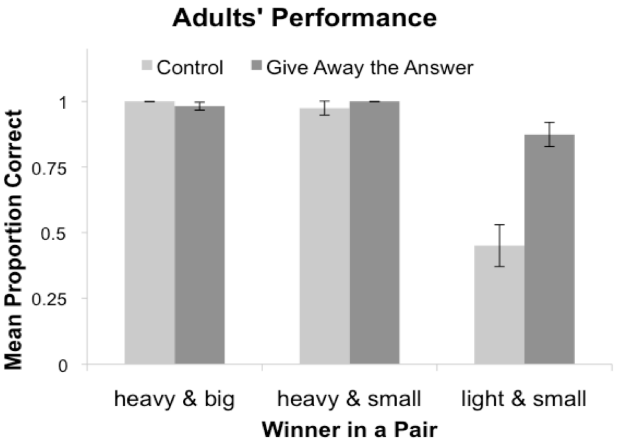

Figure 4. Mean proportion of adults’ correct choices during Experiment 2, separated by condition, and pair type. Standard errors are represented as error bars.

Adults in the control condition performed at ceiling when the faster sinking object was heavy (i.e., when the winner of a pair is heavy and big, or heavy and small). In contrast, they performed at chance when the faster sinking object was light and small – evidence of the well-documented misconception that heavy things are faster than light things. Importantly, performance improved substantially with training, $t(15) = 4.1$, $p < .01$, with adults in the *give-away-the-answer* condition performing at ceiling for all three pair types.

Children's performance shows a slightly different pattern compared to that of adults. First, children's performance success in the *control* group was low, virtually identical to children's performance in the *open-exploration* condition of Experiment 1. In other words, whether children explored sinking and floating jars prior to the assessment (Experiment 1) or not (Experiment 2), their success rate stayed the same across all three pair types ($M = .51$ to $.44$; $F < 1.2$, $p > .6$).

Second, while prediction success improved for children in the *give-away-the-answer* condition, compared to children in the *control* condition, $t(18) = 2.4$, $p < .05$, the relative improvement across pair type differs from the improvement seen in Experiment 1. In particular, in this experiment significant improvements were found when the winner was heavy and big, $t(18) = 2.4$, $p < .05$, and when the winner was heavy and small, $t(18) = 3.2$, $p < .01$, but not when the winner was light and small, $t < 1$. [1]

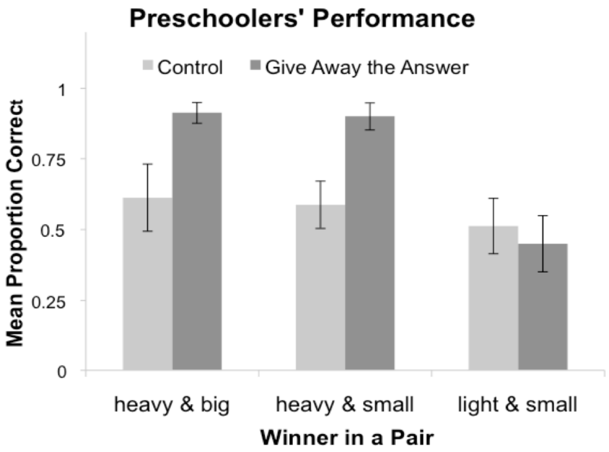

Figure 5. Mean proportion of preschoolers' correct choices during Experiment 2, separated by condition, and pair type. Standard errors are represented as error bars.

Taken together, training children to focus on the imagined empty space inside a solid object allowed them to make better predictions about sinking behavior than they would without such training. Importantly, performance improved even though participants had to abstract crucial information from the training and apply it to a new set of objects.

IV. CONCLUSION

Self-guided exploration is a common mode of learning at preschool age. Teachers of young children often shy away from direct instruction in science learning, and statewide curricula do not prescribe it. The results of the current study challenge this view. After only a short training period involving instruction, preschoolers' predictions about sinking objects improved. This was the case when the training involved the same objects that were used during testing (Experiment 1), and also when children had to extrapolate the lesson to new objects (Experiment 2). In contrast, when children were merely allowed to explore the sinking behavior of the objects on their own (*open-exploration* condition of Exp. 1), their performance was virtually equivalent to performance of children who were provided with the testing alone (*control* condition of Exp. 2).

The weak performance of students who were not given a model highlights the perils of model-free learning approaches. In other words, self-exploration on its own might not be sufficient for learning more abstract concepts such as density. Similar results have been observed in the domain of mathematics: teaching mathematical concepts directly in terms of abstract formulation of the relevant mathematical principle benefitted transfer, compared to providing students with a multitude of concrete examples of the problem, [18]. Solely providing multiple concrete examples or letting students explore the effects of physical laws on their own might lead to competition between the accidental features of the examples and the overarching regularities behind them and thus impair learning success.

Hence, we argue for the need of direct instruction focusing on relevant rules and regularities of a topic domain. Training needs to be incorporated alongside self-exploration in order for children to develop accurate understanding of a concept. This conclusion follows in line with other research on density (e.g., [19]), suggesting that instructional conversations are essential in helping children change their view of sinking and floating objects. These empirical results also have implications for the discussion about model-based versus model-free approaches for learning agents and systems in general, and developmental robotics in particular, see [9].

Of course, questions remain about the different patterns of performance we observed for children in the *give-away-the-answer* conditions of the two experiments. For example, why did training affect only performance on some types of trial and not others? And why did relative improvement fail to replicate across the two experiments? However, despite these open questions, the evidence for positive effects of training is strong. Even though children in the *give-away-the-answer* conditions might not have acquired a precise mathematical understanding of the concept of density, their learning nevertheless allowed them to move away from focusing on the surface properties of mass or volume alone – setting the stage for more advanced learning later on.

---

[1] It has to be noted that the use of a more conservative, non-parametric test (Mann-Whitney *U*) does corroborate the overall effect for adults ($z = -2.44$, $p < .01$) and preschoolers ($z = -2.04$, $p < .05$) in experiment 2, but does not reliably distinguish between the performances of children in the *give-away-the-answer* condition and the *open-exploration* condition in experiment 1 ($z = -1.13$, $p = .13$). This may be due to the comparatively low sample size, compared to the *U*-statistic's demands.

ACKNOWLEDGMENT

We thank Daniel Baum, Chris Erb, Adrienne Frazier, and Ahn Thu Inman for their help with materials, and we thank Ami Patel and Jenna Clark for their help with data collection. Data from Experiment 1 was collected as part of a research-methods class. The students who worked on this project are Shania Powell, Chanel Robinson, and Marie Roos. Funding for this paper was provided by NSF DRL 723638 and NICHD HD055324 awarded to HK).

REFERENCES

[1] E. Bonawitz *et al.*, "The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery," Cognition, in press.
[2] D. Buchsbaum, *et al.*, "Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence," Cognition, in press.
[3] American Association for the Advancement of Science, Science for all Americans: Project 2061. New York: Oxford University Press, 1990.
[4] American Association for the Advancement of Science, Benchmarks for Science Literacy. New York: Oxford University Press, 1993.
[5] National Research Council, Inquiry and the national science standards. Washington, DC: National Academy Press, 2000.
[6] P. A. Kirschner *et al.*, "Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching," Edu. Psych., vol. 41, pp. 75-86, Apr. 2006.
[7] D. Klahr and M. Nigam, "The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning," Psych. Sci., vol. 15, pp. 661-667, Oct. 2004.
[8] J. D. Bransford *et al.*, Eds., How people learn: brain, mind, experience, and school. Washington, D.C.: National Academy Press, 2000.
[9] D. Skocai *et al.*, "Formalization of different learning strategies in a continuous domain framework," Ninth International Conference on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems, Lund, Sweden, 2009.
[10] A. Mitchell and C. Savill-Smith, "The use of computer and video games for learning. A Review of the Literature," London: Learning and Skills Development Agency, 2004.
[11] C. Smith *et al.*, "On differentiation: A case study of the development of the concepts of size, weight, and density," Cognition, vol. 21, pp. 177-237, May 1985.
[12] M. Hewson and P. Hewson, "Effect of instruction using students' prior knowledge and conceptual change strategies on science learning," J. Res. Sci. Teach., vol. 20, pp. 731-743, Sep. 1983.
[13] D. Penner and D. Klahr, "The Interaction of Domain-Specific Knowledge and Domain-General Discovery Strategies: A Study with Sinking Objects," Child Develop., vol. 67, pp. 2709-2727, Dec. 1996.
[14] D. Kuhn *et al.*, The development of scientific reasoning skills. New York: Academic Press, 1988.
[15] Academic content standards, CUR GDE 420 ODE 2004 Gr.K-12, 2009.
[16] H. Kloos and G. Van Orden, "Can preschoolers' mistaken beliefs benefit learning?," Swiss J. Psych., vol. 64, pp. 195-205, Jul. 2005.
[17] H. Kloos and S. C. Somerville, "Providing impetus for conceptual change: The effect of organizing the input," Cog. Develop., vol. 16, pp. 737-759, Jun. 2001.
[18] J. A. Kaminski *et al.*, "The Advantage of Abstract Examples in Learning Math," Science, vol. 320, pp. 454-455, Apr. 2008.
[19] D. Butts *et al.*, "Is hands-on experience enough? A study of young children's views of sinking and floating objects," J. Elementary Sci. Edu., vol. 5, pp. 50-64, Feb. 1993.