Science learning in informal contexts

Behavioral studies on children's knowledge of natural phenomena and family learning in the museum

Franse, R.K.

Publication date
2021

Document Version
Other version

License
Other

2

Children's Understanding of Floating and Sinking: Predictions and Explanations Tell Different Stories

This chapter is based on: Franse, R.K., Van Schijndel, T.J.P., Visser I., & Raijmakers, M.E.J. (2020). *Children's Understanding of Floating and Sinking: Predictions and Explanations Tell Different Stories*. Manuscript in revision.

ABSTRACT

The study of children's naïve conceptual knowledge of physical principles could benefit from category learning research. The present study systematically investigated the development of children's knowledge of buoyancy, by considering learning principles revealed by this research. Four- to twelve-year-old children (N = 139) were asked to predict buoyancy for different types of objects and to provide explanations for their answers. Three sets of objects differing in appearance were presented: cubes with indistinctive appearance, differing in density, volume and mass (Set 1), cubes made from a distinctive material, wood and metal (Set 2), and well-known objects that are exemplar floaters or sinkers, e.g. a coin and a boat (Set 3). To account for individual differences in knowledge representation, responses on the prediction task were modeled with latent regression analysis. Results showed that children integrated mass and volume in predicting buoyancy of Indistinctive cubes (Set 1), and integrated mass, volume and material for predicting buoyancy of Material cubes (Set 2). This integrated information remained largely implicit because very few children gave explicit explanations of buoyancy that agreed with their predictions. Explanations about buoyancy mostly agreed with simple rules (for Set 1 and 2) or simple facts about exemplars (Set 3). Solution strategies in predictions (but not the explanations) showed a clear developmental pattern, such that the way mass and volume were integrated improved with age.

Keywords: conceptual knowledge, naive physics, proportional reasoning, multiple-process theories, cognitive development

2.1 INTRODUCTION

Laymen's knowledge of lawful, physical phenomena, such as gravity, the torque principle, shadow size, and buoyancy is based on heterogeneous sources of information. Buoyancy clearly illustrates the heterogeneity of learning contexts: bathing generates perceptual experiences, coins thrown into a fountain and boats on a river are distinct exemplars children observe, and Archimedes' law is referred to in informal and formal science education for children. The type of conceptual knowledge resulting from this kind of heterogeneous information is of great interest for cognitive science (Machery, 2010), and students' conceptual knowledge of physical concepts is also of societal relevance as one of the key elements of scientific literacy (Vayssettes, 2016).

2.1.1 Children's Understanding of Physical Phenomena

Conceptual knowledge of a physical phenomenon is defined as abstract, general principles about the phenomenon, which can either be explicit, such that it can be expressed in words, or implicit, such that it is not verbalizable (Rittle-Johnson, Schneider, & Star, 2015). Evidence for different types of naïve conceptual knowledge has been reported in the literature for an abundance of physical phenomena. In some cases, to show that representations change with development (e.g., Edelsbrunner, Schalk, Schumacher, & Stern, 2015; Flaig et al., 2018; Hardy, Jonen, Möller, & Stern, 2006; Schneider & Hardy, 2013) or evolve depending on the learning environment (Hofman, Visser, Jansen, & Van der Maas, 2015), but also in the context of the scientific dispute about the existence of mental models versus fragmented knowledge (DiSessa, Gillespie, & Esterly, 2004; Vosniadou, 2008). One could distinguish three different types of naïve knowledge. First, there is ample evidence for verbalizable coherent, but often incomplete or even erroneous conceptual knowledge for various domains: Buoyancy (Flaig et al., 2018; Schneider & Hardy, 2013), the torque principle (Jansen & Van der Maas, 1997; Siegler, Strauss, & Levin, 1981), shadow size (Van Schijndel & Raijmakers, 2016), object motion (Hast & Howe, 2012); see Duit (2009) for an overview of students' potential misconceptions. These, often called, misconceptions consist of simple rules: a one-dimensional principle or two dimensions that are combined in an additive way. With development, these rules become increasingly complex and approximate reality better (Siegler et al., 1981). Second, there is also evidence for naïve conceptual knowledge that consists of principles for genuine information integration (i.e., not additive but multiplicative), such as in the domains of the torque principle (Hofman et al., 2015; Wilkening & Anderson, 1982; Wilkening & Cacchione, 2011), buoyancy (Janke, 1995), shadow size (Ebersbach & Resing, 2007), and velocity, time and distance (Wilkening, 1981). An important developmental insight is that the integration of information improves with age, that is, a bias in weighting the dominant dimension decreases (Ebersbach & Resing, 2007; Janke, 1995; Wilkening & Cacchione, 2011). However, the use of explicit information about the situation may temporarily hinder information integration (Strauss &

Stavy, 1982; Wilkening & Cacchione, 2011). Third, there is evidence for the understanding of physical phenomena that consists of exemplars taken from everyday observations, which does not form an abstract principle, but is still generalizable to similar situations. Such fragmented knowledge has been reported in the domains of buoyancy (Schneider & Hardy, 2013), the earth (Straatemeier, Van der Maas, & Jansen, 2008) and force (DiSessa et al., 2004).

Interestingly, these different types of conceptual knowledge of physical phenomena agree quite well with different types of knowledge representations that are reported in more fundamental, category learning studies, that is, studies with artificial instead of natural categories (Ashby & Ell, 2001; Erickson & Kruschke, 2002).

2.1.2 Insights from Human Category Learning

One of the central questions in category learning research concerns the knowledge people construct of a categorization that they get acquainted with from classifying or observing examples (Ashby & Maddox, 2005; Kruschke, 2005). Category learning research has been revealing several learning principles that seem relevant for studying the development of naïve conceptual knowledge of physical principles. First, people use different solution strategies to solve more complex categorization tasks: rule-based, information integration, and exemplar solution strategies (Ashby & Gott, 1988; Johansen & Palmeri, 2002; Raijmakers, Schmittmann & Visser, 2014; Wills, Inkster, & Milton, 2015). Physical principles, such as force, the torque principle, and buoyancy, typically generate difficult categorization problems, such as floating versus sinking. After all, physical principles integrate multiple dimensions with a multiplication or quotient relation and these principles are typically difficult to discover from everyday observations (Jansen & Van der Maas, 1997; Schultz & Takane, 2007; Siegler et al., 1981). Hence, it is expected that in the development of conceptual knowledge about these principles, people apply suboptimal solution strategies that differ within and between people.

Second, the type of solution strategies people develop during a categorization task depends on the category structure to be learned: rule-based, information-integration, or other complex category structures, such as rules with exceptions (Ashby & Ell, 2001; Erickson & Kruschke, 1998). Many physical principles are not only integrating multiple dimensions, everyday observations make it even more difficult to learn them. These observations do not only differ on a few relevant dimensions, but distinct exemplars are part of everyday observations too. For example, boats on the river are distinct exemplars with physical features, a heavy weight and built from metal with a distinct shape, that are difficult to integrate in an abstract principle about floating and sinking that works for other objects, such as stones and coins. Hence, it is expected that depending on the way a physical phenomenon is presented, different representations of naïve conceptual knowledge will be revealed.

Third, from category learning research has become clear that behavioral measurements not always reveal the same solution strategy as verbal explanations. Information integration representations are mostly evidenced in behavioral data, that is, the prediction of category membership of unseen examples. Usually, verbal explanations of these predictions do not agree well with the behavioral data (Ashby & Waldron, 1999). Studies about conceptual knowledge of physical principles differ in the type of measurements that have been applied (Straatemeier et al., 2008; Van Schijndel, Van Es, Franse, Van Bers, & Raijmakers, 2018b). When both predictions and explanations are assessed, outcomes do not always agree (Pine & Messer, 1999). Hence, to reveal knowledge representations it is of interest to test whether verbal explanations agree with solution strategies as appearing from behavioral, prediction data.

Buoyancy, floating and sinking, is a typical phenomenon with complex physical principles and distinct exemplars in daily life, that is, a complex domain to learn from numerous, informal experiences. Buoyancy is discussed in primary school in many countries, but in the Netherlands formal principles (density and the Law of Archimedes) are only introduced in Grade 7, secondary school. In the current study we aim to reveal conceptual knowledge about buoyancy in primary school children. Considering what we learned from categorization studies, we will introduce multiple stimulus sets, which can be classified in floaters and sinkers in different ways, we will test for the use of latent solution strategies, and we will assess both predictions and explanations. We expect children to express different types of conceptual knowledge: rule-based (RB) representations, information-integrating (InI) representations, and exemplar-based (EB) representations, depending on type of objects we present and the measurement methods we use.

2.1.3 Previous Studies on Floating and Sinking

When explaining buoyancy in interviews, children refer to a range of characteristics (mass, volume, containing air, materials), but the explanation referring to mass as the only dimension is the most commonly used (Butts, Hofman, & Anderson, 1993; Hsin & Wu, 2011; Smith, Carey, & Wiser, 1985; Smith, Maclin, Grosslight, & Davis, 1997; Tenenbaum, Rappolt-Schlichtmann, & Zanger, 2004). A large systematic study on buoyancy with children judging verbal statements (Edelsbrunner et al., 2015; Schneider & Hardy, 2013) suggests that different mental models are prevalent at different ages. Schneider and Hardy distinguished three knowledge profiles: (a) misconceptions (one-dimensional focus on mass, volume, or shape, or considering air an active force that pulls objects), (b) everyday conceptions (considering the role of water, the concept of material kind, and the hollowness of objects as explanations of buoyancy), and (c) scientific concepts[4] (considering one or more of the physical quantities of density, water pressure, or buoyancy force). These results agree with rule-based knowledge representations as proposed by Inhelder and Piaget (1958) that a conceptual understanding of buoyancy based on only one-dimension is prevalent in young children, before education (see also Schneider & Hardy, 2013). In contrast, the concept of buoyancy is also studied using

[4] Archimedes' Principle states that an object immersed in a fluid, is buoyed up by a force equal to the weight of the fluid displaced by the object. If the density of an object is greater than that of the fluid in which it is immersed, it sinks. If the object is either less dense than the fluid or is shaped appropriately (as a boat), the upward force of buoyancy can keep the object afloat. The density of water is 1 g/cm^3

explanation measurements. These studies reveal a different type of conceptual knowledge. At preschool age, children are already able to predict correctly whether a block sinks or floats when placed in water (Janke, 1995; Kloos, Fisher, & Van Orden, 2010; Kohn, 1993; Leuchter, Saalbach, & Hardy, 2014). Tested in this way, children show to have some sense of density, integrating mass and volume (Kloos et al., 2010; Kohn, 1993). The study of Janke (1995) shows evidence that children integrate information about mass and volume of boats in predicting their buoyancy and thus that children have information-integration representations of buoyancy. In sum, the literature offers ample evidence for the existence of multiple representations of naïve conceptual knowledge of buoyancy. However, factors that play a role in the measurement of knowledge representations were never varied systematically for buoyancy. This is the aim of the current study.

2.1.4 Current Study

The aim of this paper is to reveal how knowledge about buoyancy is represented in children before they have attained formal education in this domain (RQ-1), and how such representations develop (RQ-2). To this end, we tested 139 Dutch primary school children of different ages (4 to 12 years). We used two different tasks, a prediction task and an explanation task (structured interview). Both tasks include three types of 3D objects differing in appearance: Set 1 (Indistinctive cubes): cubes with different density, volume and mass, with equal, indistinctive appearance; Set 2 (Material cubes): cubes made from distinctive material, that is, wood and metal; Set 3 (Exemplar objects): objects that are exemplar floaters and sinkers, for example, a coin. Taking previous literature into account, we consider rule-based representations to be indicated by: a) a one or multiple dimensional solution strategy in predictions; and b) a similar solution strategy in explanations (for the same items); We consider information-integration representations to be indicated by: a) a solution strategy integrating multiple dimensions in predictions and b) a dissimilar solution strategy in explanation of buoyancy of the same objects; We consider exemplar-based representations to be indicated by a) explanations that refer mainly to similar, typical exemplars and b) which might go together with high accuracy predictions if the agreement between object and exemplar is relevant.

Our expectations are that:

1) Children's buoyancy predictions of indistinctive cubes (Set 1) reveal information-integration representations, implying that children base their predictions, but not their explanations, on an integration of multiple object dimensions (mass and volume).
2) Children's buoyancy predictions of material cubes (Set 2) reveal rule-based representations, implying that they use the same solution strategy (e.g., 1D material–rule, or a multiple dimension rule) for predictions and explanations.

3) Children’s buoyancy predictions of exemplar floaters and sinkers (Set 3) reveal exemplar-based representations, indicated by highly accurate predictions (more accurate than comparable objects of Set 2), and by the mentioning of facts or past experiences in explanations.
4) Children's predictions become more advanced with age, such that the integration of mass and volume becomes more accurate (Janke, 1995; Wilkening & Anderson, 1982). Children’s explanations become more advanced with age, that is, more scientific concepts and less misconceptions are mentioned (Schneider & Hardy, 2013).

2.2 METHOD

2.2.1 Participants

In this study, 139 children (76 boys, age: $M = 8.42$, $SD = 2.42$) from a Dutch primary school participated. One further child entered the test room, but withdrew before joining the procedure. An active consent procedure was used, with parents being required to sign and return a form if they agreed to let their child participate. The sample consisted of children between 4 and 12 years old (see Table 2.1).

Table 2.1
Age and gender distribution of participants.

Age (years)	4	5	6	7	8	9	10	11	12
Boys	5	8	9	7	8	11	12	8	8
Girls	8	9	9	7	10	5	6	6	3
Total (*N*)	13	17	18	14	18	16	18	14	11

2.2.2 Materials

Materials consisted of three sets of objects. The first set of objects consisted of cubes that could only be categorized correctly in floaters and sinkers by integrating the perceptible physical characteristics volume and mass (Set 1, Indistinctive cubes). The second set of objects consisted of wooden and metal cubes could also be categorized by a simple material rule (Set 2, Material cubes). The third set of objects could be categorized by using common facts about the buoyancy of exemplar objects, such as, coins at the bottom of a fountain (Set 3, Exemplar objects). In addition to the objects in these three sets, two reference cubes were added to illustrate the meaning of floating (object A) and sinking (object B) and to give children the option to compare dimensional values (volume and mass). Below we describe the objects in more detail. In total there were 24 objects, a plastic aquarium (size: l = 34 cm, w = 20 cm, h = 22 cm) filled with water, a drawing of a floating cube, a drawing of a sinking cube, six plastic baskets to store objects.

Indistinctive cubes (Set 1, 8 objects). Eight solid cubes differing in density, volume and mass, were made from four different colors modeling clay (Plasticine), and (hidden) polystyrene foam (Styrofoam) and/or metal bullets. To provide the cubes an equal, indistinctive appearance, the surfaces of all cubes were of the same material (i.e., modeling clay). These indistinctive cubes (see also Table 2.2) were designed to meet two requirements. First, using the reference cubes, children should be able to predict buoyancy (correctly or incorrectly) unambiguously from the perspective of three solution strategies to predict buoyancy: 1) an information-integrating, density strategy, 2) a one-dimensional volume strategy and 3) a one-dimensional mass strategy. For this reason, the value of density, volume and mass of Set 1 cubes is either lower than, or equal to, reference cube A (0.7 g/cm^3, 216 cm^3, 151 g) or higher than, or equal to, reference cube B (1.2 g/cm^3, 343 cm^3, 421 g). Second, since children's strategies will be determined based on prediction responses, the combination of objects should provide a distinct response patterns for each expected strategy. To distinguish the volume from the mass strategy objects B and H were informative (see Table 2.2). To distinguish the density strategy from the volume, or mass strategy objects D, E, F and G were informative. For example: object D has a density (0.7 g/cm3) equal to reference cube A (floater), however, the volume (729 cm^3) and mass (510 g) of object D is higher than reference cube B (sinker). Therefore, from a density strategy perspective object B would be classified as floater, while from a volume or mass strategy perspective object B would be classified as sinker. Note that solid cubes with a density higher than the density of water (1 g/cm^3) sank. The four colors were related to four different densities to assure that an explanation based on material could be consistent, avoiding fanciful theories about the insides of the cubes. To control for pre-existing associations between buoyancy and color two versions of Set 1 objects were constructed with different color-density combinations.

Air cubes (Set 1b, 3 objects). These cubes contained one or two visible holes to trigger a possible air strategy. Cube A (floater, edge length 7 cm) was a hollow bin. Cubes B (sinker, edge length 7 cm) and C (floater, edge length 6 cm) had a hole, straight through the cube. The cubes were also made from modeling clay with polystyrene foam and/or metal bullets inside. When handling over the cubes to the child, the hole was located at the upper side. As it appeared, the accuracy of the Set 1b cubes (M = .56, SD = .287), was not above chance level, were not mutually related, and only marginally related to age, $F(1,136)$ = 3.475, p = .064. Hence, these cubes (Set 1b) were not analyzed further nor were they included in the strategy analysis. Hence, they are not mentioned in the Results section.

Material cubes (Set 2, 6 objects). Three wooden cubes made of solid pine, had a density of 0.58 g/cm3 and differed in volume (edge lengths 3-8 cm). Three metal cubes were made of solid aluminum, had a density of 2.8 g/cm3 and differed in volume (edge length 2-6 cm).

Exemplar objects (Set 3, 5 objects). The five exemplar floaters and sinkers were a toy boat, a wood stump, a 50-euro cent coin, a ball, and a pebble. Based on expected solution-strategies for Set 1 predictions (i.e., density strategy, volume strategy, mass strategy), we included a small, light object with a high-density that sinks (coin) and a large, heavy object with a low-density that floats (wood stump). To find out whether the children recognized the objects, each time a new object was introduced, the children were asked to name it aloud. All children succeeded in doing so.

2.2.3 Procedure

Children were tested individually by one of two female experimenters in a private room at their school. The child was facing the experimenter at a table. The experimental procedure consisted of two phases, a prediction phase followed by an explanation phase. In the prediction phase the objects were presented – one by one – in three sets: Set 1, Set 2, and Set 3. On the table an aquarium with water was visible for the child throughout the task. There were two sequences of presentation within sets (one being the reversed order of the other) and one fixed sequence of presentation between sets. The experimenter used the reference cubes to introduce the task and the concept of floating and sinking. At the beginning of each set, the child tested the buoyancy of the reference cubes by putting them into the water. Then, the child was asked to place the reference cubes either on the left side of the table next to a floating object picture, or on the right side of the table next to a sinking object picture. If the child was able to do this correctly, the prediction task started. The experimenter handed over the first object of Set 1, and asked the child to predict float or sink. The child was allowed to compare the object's volume and mass with the reference cubes, by lifting and holding the object and reference cubes. After making a prediction the child was asked to place the object next to the floating object picture or the sinking object picture. While handing over the next object to the child, the experimenter put the previous predicted object into one of two baskets. After the child had predicted buoyancy of the last object of Set 1, the experimenter took the two baskets with predicted objects off the table and placed two empty baskets. The procedure of testing reference cubes and predicting objects continued till buoyancy was predicted of 22 objects.

In the explanation phase the child was asked to justify his or her predictions for each set of objects separately during a structured interview. To this end, the child was presented with two baskets that contained the objects from one set as sorted by the child into float and sink. The open-ended questions asked during the interview were: "*About the objects in this basket* [pointing at the basket 'float'] *you said that they will float if you put them in water. About the objects in this basket* [pointing at the basket 'sink'] *you said that they will sink if you put them in water. Can you tell me why?"* After the child had the opportunity to respond, the additional question *"Can you explain it to me a little more?"* was asked once. The interviews were recorded on video, and final scoring was based on transcripts (in CLAN; MacWhinney & Snow, 1990) of the videotapes.

Table 2.2
Characteristics of test objects and accuracy of predicting buoyancy per object.

Name	Object Properties					Object Sequence		Object Difficulty
	Density	Length	Volume	Mass	Buoyancy	Order 1	Order 2	Prop.
	(g/cm^3)	(cm)	(cm^3)	(g)				(*N*)
Reference cubes								
A	0.7	6	216	151	Floats	1	1	
B	1.2	7	343	421	Sinks	3	3	
Indistinctive cubes (Set 1)								
A	0.3	2	8	3	Floats	7	7	.82 (114)
B	0.3	7	343	103	Floats	10	4	.81 (112)
C	0.7	2	8	6	Floats	2	2	.65 (91)
D	0.7	9	729	510	Floats	5	9	.29 (41)
E	1.2	4	64	77	Sinks	4	10	.35 (48)
F	1.2	5	125	150	Sinks	8	6	.37 (51)
G	4.5	3	27	121	Sinks	6	8	.85 (118)
H	4.5	5	125	562	Sinks	9	5	.93 (129)
Air cubes (Set 1b)								
A	-	7	-	232	Floats	11	13	.47 (66)
B	-	7	-	234	Sinks	12	12	.71 (98)
C	-	6	-	138	Floats	13	11	.50 (70)
Material cubes (Set 2)								
A	0.58	3	27	16	Floats	18	15	.83 (116)
B	0.58	6	216	125	Floats	16	17	.81 (112)
C	0.58	8	512	297	Floats	15	18	.77 (107)
D	2.8	2	8	22	Sinks	19	14	.78 (108)
E	2.8	3	27	76	Sinks	14	19	.91 (126)
F	2.8	6	216	605	Sinks	17	16	.91 (127)
Exemplar objects (Set 3)								
Boat	l = 26, w = 10, h = 3				Floats	20	24	.89 (124)
Wood	h = 18, d = 7				Floats	21	23	.51 (71)
Coin	h = 0.2, d = 2.5				Sinks	22	22	.91 (126)
Ball	d = 11				Floats	23	21	.88 (123)
Pebble	l = 4, w = 3, h = 3				Sinks	24	20	.92 (128)

Note. Listed object properties are density, length (edge length of cube), volume, mass, and buoyancy. The sizes of the exemplar objects are given in l = length (cm), w = width (cm), h = height (cm), d = diameter (cm). The reported accuracy is the proportion (*N*) of children correctly predicting buoyancy for the object.

2.2.4 Analysis

Explanations provided by the children were analyzed by scoring different wording used in them; for Set 1 they were scored into five categories: 1) scientific, 2) mass and volume, 3) only volume, 4) only mass, and 5) residual. For sets 2 and 3 additional scoring categories were used to reflect knowledge of either specific materials (Set 2) or specific facts about objects (Set 3).

The analysis of the prediction data is able to capture one-dimensional (e.g., mass-strategy, material-strategy) and multiple-dimensional (e.g., density strategy, mass-material strategy) solution strategies and to capture heterogeneity. We used a hybrid model based on latent class regression analysis (Huang & Bandeen-Roche, 2004). See e.g. Bouwmeester, Sijtsma, & Vermunt (2004) and Hofman et al. in (2015) for similar models applied in cognitive development.

In the latent class regression model, heterogeneity is captured by multiple classes, which are each characterized by a distinct pattern of prediction responses indicative of a particular strategy. For example, one class may have children that only use the mass of objects to predict buoyancy whereas another class of children combines both mass and volume. The latent class regression model captures this by having multiple classes of participants with different weighted combinations of mass, volume, and other features of the objects determining the response. In particular, the model is defined as follows:

$$\log(p(y_i \mid c/(1\text{-}p(y_i \mid c)) = \alpha_c + \beta_{mass,c}\log(\text{mass}) + \beta_{volume,c}\ \log(\text{edge}) + \beta_{color,c}\,\text{color}$$
$$c = 1 \ldots N_c,\ i = 1 \ldots n$$

Here y_i is the response (either 'float' or 'sink'); $p(y_i)$ is the probability of such a response in the model, which is conditional on latent class c; parameters $\beta_{mass,c}$, $\beta_{volume,c}$, $\beta_{color,c}$ are the regression coefficients for latent class c; N_c is the number of latent classes; and n is the number of participants. Hence, the probability of a 'sink' response is determined by a weighted combination of the volume and mass of the object, and possibly other features such as color in above equation. Note that the logarithms of mass and volume are used as the perception of mass and volume is better described on a logarithmic scale, i.e. a Fechner scale, than a linear scale (Jones, 1986; Murray, 1993).

To see how this model can capture response patterns based on either RB or InI knowledge representations of buoyancy, consider the following examples. An expected, rule-based knowledge representation of floating and sinking is based mass only. The model can capture such a strategy by setting $\beta mass > 0$, $\beta volume = 0$, $\beta color = 0$, $and\ a_c > 0$. That is, children only use the mass of objects to determine their response, and ignore volume and color of the objects. Setting both $\beta mass$ and $\beta volume$ larger than zero results in a strategy that integrates both these dimensions and hence is consistent with either a multi-dimensional rule-based representation (if predictions and explanations are consistent), or an information-integration-based representation (if predictions and explanations are inconsistent).

To arrive at a parsimonious description of the prediction performance data, latent class regression models including parameters a_c, β_{mass} and β_{volume} (henceforth, base models) are fitted to the prediction data with different numbers of classes, in particular 1-3 class models are fitted to limit the total number of parameters. To select the optimal number of classes we used the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC; Schwarz, 1978) by choosing the models with the lowest values on these criteria. After selecting the number of classes we simplified the optimal model in several ways and we selected, what we call the best fitting, but most parsimonious model (doing justice to the trade-of between fit and parsimoniousness) as indicated by the AIC and or the BIC. It is not uncommon that the AIC and BIC point to different models, in which case we compare simplifications of models with different number of classes (see also Schneider & Hardy, 2013).

The final step in the analysis of prediction performance data is to classify individual performance. Using the latent class regression model, this is achieved by computing the so-called posterior probabilities, i.e., the probability of membership of the classes of the model conditional on a particular pattern of responses. Children were assigned to the class with the highest posterior probability of their responses for each set of objects separately (Visser, 2011).

2.3 RESULTS

This section is divided into three parts. First, we present the results for children's predictions of buoyancy, which includes the analysis of solution strategies for each object set separately and a comparison between sets of objects. Second, we present the results for children's explanations. The third part of the results is focused on the consistency between predictions and explanations, which can shed light on possible different type of knowledge representations (i.e., rule-based (RB), information-integration (InI), and exemplar-based (EB).

2.3.1 Prediction Phase

General results. Comparing the prediction performance between sets of objects, a repeated measures ANCOVA on the proportion correct with object sets (Set1, Set2, Set3) as within subject factor, gender as between-subject factor and age as covariate showed a significant main effect of objects, $F(1.93,262.47) = 9.75$, $p < .01$ and a significant contribution of age, $F(1,136) = 69.31$, $p < 0.01$ (with corrected degrees of freedom using Huynh-Feldt estimates of sphericity, $\varepsilon = 0.97$ for the main effect of objects). No significant contribution of gender was found, nor any two-way interactions. Contrasts revealed that the accuracy of Set 2 (material cubes; $M = .83$, $SD = .222$), $F(1,136) = 6.26$, r = 0.21, and Set 3 (exemplar objects; $M = .82$, $SD = .202$), $F(1,136) = 20.81$, r = 0.36 was significantly higher than the accuracy of Set 1 (indistinctive cubes, $M = .63$, $SD = .206$). Accuracy (proportion correct) of all three sets was significantly above the 50% probability level with $t(138) = 7.621$, $p < .001$ (Set 1), $t(138) = 17.733$, $p < .001$ (Set 2), and $t(138) = 18.895$, $p < .001$ (Set 3), respectively.

Strategy analysis. To reveal the presence of different strategies in predicting buoyancy, we applied latent class regression analysis to model the probability of a prediction (floating vs. sinking) as a function of cube characteristics. In the base models for Set 1 and Set 2 log(mass) and log(edge length) were included as predictors. Extended models included color (Set 1) or cube material (Set 2) as predictors. Since the exemplar objects of Set 3 had no systematic variation of volume, mass or other characteristics, the predictions for the five objects (boat, wood stump, coin, ball and pebble) were analyzed in a latent class model without additional predictors.

Indistinctive cubes (Set 1). From the fitted base models the 2-class (BIC) and 3-class (AIC) models are the best fitting, most parsimonious models (see Table S2.1 in 2.6 Supplementary material). Additional 2-class and 3-class models were fitted by successively setting prediction parameters to zero in order to find a more parsimonious model. This resulted in a three-class model with two classes defined by mass and volume (Strategy 1a, $\alpha = 102$, $\beta_{mass} = -31.0$, $\beta_{volume} = 14$ and Strategy 1b, $\alpha = 3.7$, $\beta_{mass} = -3.0$, $\beta_{volume} = 2.2$), and one class defined by color (Strategy 1c, $\alpha = -0.3$, $\beta_{color1} = -1.1$, $\beta_{color2} = 0.6$, $\beta_{color3} = 1.0$). This model is represented graphically in Figure 2.1.

In Figure 2.1, Strategy 1a (red curves and stars), Strategy 1b (green curves and stars) and Strategy 1c (blue stars) are depicted in multiple graphs (volume = 8, volume = 27, volume = 64, volume = 125, volume = 343, volume = 729 cm^3). The vertical black line represents the density of water (1 g/cm^3): cubes left of this line actually float, and cubes right of this line sink. Each graph shows the by the model predicted probability of 'predicting floating' as a function of mass at a constant volume. In the fourth graph (volume 125), for example, the logistic curve of Strategy 1a shows that cubes with a mass less than 125 gram were predicted by the model as having a high probability to float. Cubes with a mass greater than 125 grams were predicted by the model as having a low probability to float. A (hypothetical) cube with a volume of 125 cm^3 weighing 125 grams, and therefore with a density equal to water (1.0 g/cm^3), was predicted by the model with a 50% chance of floating, which was the most rational outcome if one is forced to choose between floating and sinking. Hence, children with Strategy 1a integrated mass and volume properly for cubes with edge 5 cm (i.e., volume 125 cm^3). However, integration of volume and mass was biased towards floating for smaller cubes with density close to 1 g/cm^3. In the corresponding graphs (volume = 64, volume = 27, volume = 8) the curve and vertical line (water) do not intersect at a probability of 50%. Instead, the curve is shifted to the right. For small volumes children *under*estimated the cube's mass if its density was close to one, and they would predict that a (hypothetical) cube with volume 8 cm^3 and a mass of 15 grams (i.e. a density of 1.9 g/cm^3) to float. For very large cubes (volume 729 cm^3) an opposite effect is visible. The curve and vertical line intersect at a higher mass, and the curve is shifted to the left. According to the model, children *over*estimated the mass, and mistakenly predicted that a cube with a volume of 729 cm^3 and a mass of 510 grams (i.e. a density of 0.7 g/cm^3) would sink. According to the model, children applying Strategy 1b, also integrated the dimensions mass and volume, as can be seen in the logistic curve (green curve). For this strategy the *under*estimation of mass in smaller volumes was much more extreme than in Strategy 1a. Strategy 1c was based on the cube color, which is a nominal variable. For this reason, this strategy is depicted in Figure 2.1 by separate points (blue stars). We interpret Strategy 1c as a residual strategy because it shows much more uncertain responses, that is, P(floating) is close to .5 for several objects. Moreover, the relation between colors and buoyancy predictions is not consistent. Strategy use is related to age (Table 2.3, Figure 2.2a) as shown by logistic regression (Cox and Snell $R^2 = .297$, medium effect) for Strategy 1a ($\beta_{age} = .725$, $p < .001$) and for Strategy 1b ($\beta_{age} = .482$, $p < .001$), relative to the residual strategy.

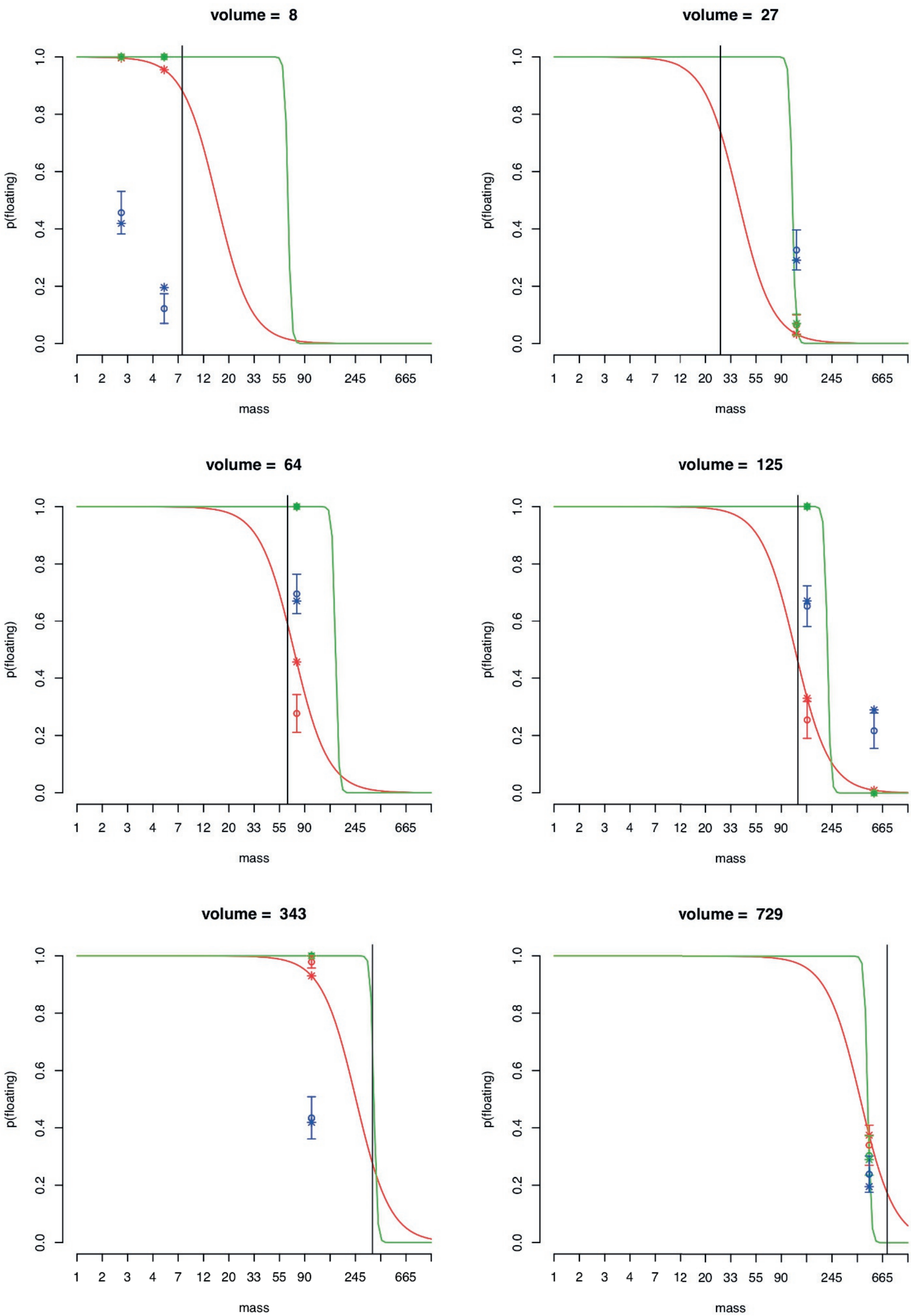

volume = 8
volume = 27
volume = 64
volume = 125
volume = 343
volume = 729
p(floating)
mass
0.0 0.2 0.4 0.6 0.8 1.0
1 2 3 4 7 12 20 33 55 90 245 665

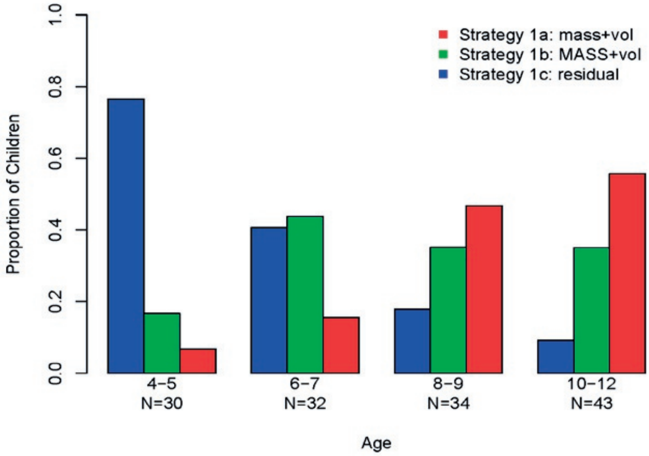
Figure 2.1. Indistinctive cubes (Set 1): Three strategies as described by the optimal model of prediction responses for Set 1. Each modeled strategy is depicted by multiple figures for constant cube volume as a function of mass: volume = 8, volume = 27, volume = 64, volume = 125, volume = 343, volume = 729 cm^3 (corresponding to cube edge lengths of 2 cm, 3 cm, 4 cm, 5 cm, 7 cm and 9 cm respectively). Each graph shows the modeled strategies (curves and stars) and observed values (circles with standard errors) for prediction responses. Strategies are depicted with different colors: red curves, circles and stars for Strategy 1a, green curves, circles and stars for Strategy 1b, blue circles and stars for Strategy 1c. Note that some standard errors are close to zero, moreover, model and observations might overlap completely (especially for the green data points). The vertical black line represents the density of water (1 g/cm^3): cubes left of this line actually float, and cubes right of this line sink. See text for further details. Note that two graphs show observed values for two objects (i.e., volume = 8, cube A and C; volume = 125, cube F and H) and four graphs show observed values for one object (i.e., volume = 27, cube G; volume = 64, cube E; volume = 343, cube B; volume = 729, cube D).

Table 2.3
Characteristics of solution strategies for the three types of objects.

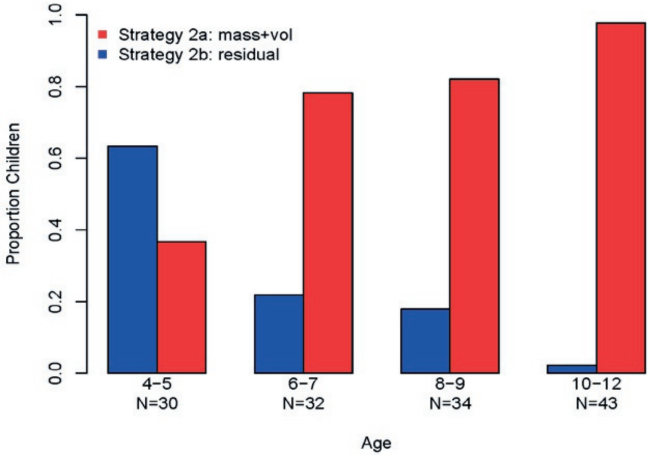
Strategy (characteristic)	Accuracy (*SD*)	Age in yr (*SD*)	% (*N*)
Indistinctive cubes (Set 1)			
Strategy 1a (mass+vol)	.83 (.09)	9.5 (1.87)	34 (47)
Strategy 1b (MASS+vol)	.65 (.08)	8.2 (2.08)	33 (46)
Strategy 1c (residual)	.42 (.17)	6.1 (2.09)	33 (46)
Material cubes (Set 2)			
Strategy 2a (mass+vol)	.93 (.11)	8.6 (2.23)	76 (106)
Strategy 2b (residual)	.53 (.21)	5.8 (1.92)	24 (33)
Exemplar objects (Set 3)			
Strategy 3a (facts)	.89 (0.12)	8.3 (2.42)	86 (120)
Strategy 3b (residual)	.42 (0.15)	5.7 (1.37)	14 (19)

Note. Accuracy for predicting buoyancy, the average (standard error) of age, and the percentage of children per strategy for Indistinctive cubes (Set1), Material cubes (Set 2) and Exemplar objects (Set 3). Strategy names are explained in the text.

(a)



(b)



(c)

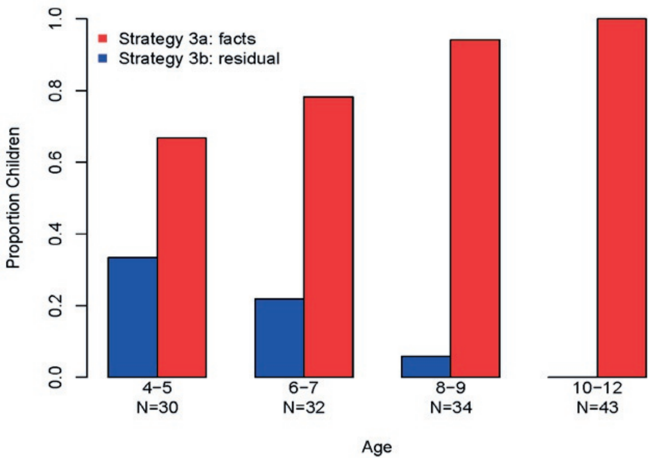


Figure 2.2. Proportion of children using a strategy for making predictions about buoyancy within each age group for respectively a) Indistinctive cubes (Set 1), b) Material cubes (Set 2), and c) Exemplar objects (Set 3).

Material cubes (Set 2). We followed the same modeling procedure as for the Set 1 objects and fitted latent class regression models with 1 through 3 classes. Testing a series of models (see Methods, and see Table S2.2 in 2.6 Supplementary material) showed that the base model was optimal with 2 classes (based on both AIC and BIC). Hence, we tested more parsimonious 2-class models in order to find the most optimal model. This resulted (based on the BIC) in a two-class model with one class defined by mass and volume (76% of the children) in addition to an intercept (Strategy 2a, , $\alpha = 7.98$, $\beta_{mass} = -6.23$, $\beta_{volume} = 4.66$) and one class defined by an intercept only, i.e. a constant probability floating (24%; Strategy 2b, $\alpha = -0.48$). The AIC favors a 2-class model that includes a strategy based on mass, volume and cube material in addition to a residual strategy (see Table S2.2 in 2.6 Supplementary material). We will show the more parsimonious model here.

The model is shown graphically in Figure 2.3. In the third graph (volume 216) the logistic curve of Strategy 2a (red curve) shows that for middle sized cubes buoyancy is correctly estimated. As with Strategy 1a (Set 1), we see that for smaller cubes, the Strategy 2a curve is shifted to the right, relative to the black vertical line (the line that shows the density of water). This indicates that, according to the model, children underestimated mass for small volumes. For large cubes the curve is shifted to the left, which means that children overestimate the mass of the cube. Indeed, relatively many prediction errors were made for the smallest metal cube and the largest wooden cube. Strategy 2b is not systematically related to mass or volume. Strategy use is related to age (Table 2.3, Figure 2.2b) as shown by logistic regression (Cox and Snell $R^2 = .233$; $\beta_{age} = .612$, $p = .001$) relatively to the residual strategy.

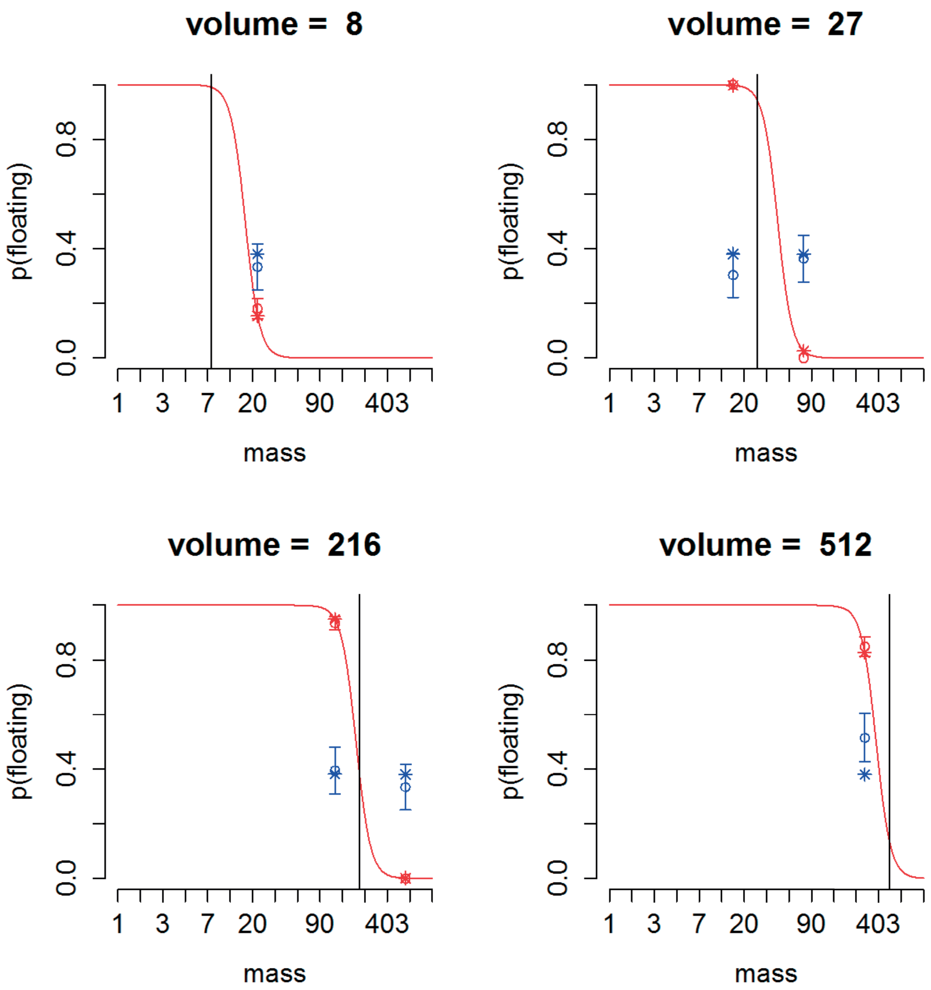


Figure 2.3. Material cubes (Set 2): Two strategies as described by the optimal model of prediction responses for Set 2. Each modeled strategy is depicted by multiple figures for constant cube volume as a function of mass: volume = 8, volume = 27, volume = 216, volume = 512 (corresponding to cube edge lengths of 2 cm, 3 cm, 6 cm, 8 cm respectively). Each graph shows the modeled strategies (curves and stars) and observed values (circles with standard errors) for prediction responses. The red curves, circles and stars relate to Strategy 2a. The blue circles and stars relate to Strategy 2b. The vertical black line represents the density of water (1 g/cm^3): cubes left of this line float, and cubes right of this line sink. See text for further explanation. Note that two graphs show observed values for two objects (i.e., volume = 27, cube A and E, and volume = 216, cube B and F) and two graphs show observed values for one object (i.e., volume = 8, cube D, and volume = 512, cube C).

Exemplar objects (Set 3). Set 3 consists of 5 objects (boat, wood stump, coin, ball, pebble) that do not vary systematically in mass and volume. Hence, we fitted to the data Latent Class Models instead of Latent Class Regression models. In addition, to test whether age would better than multiple strategies explain relations between objects, we fitted a one-class regression model with age as a covariate on the response parameter for each object. Fit statistics (see Table S2.3 in 2.6 Supplementary material) revealed that a model with two classes is the most optimal model (based on both BIC and AIC). One class (86% of the children, Strategy 3a) shows high scores for all objects (P_f(boat) = .95, P_f(coin) = .99, P_f(ball) = .96, P_f(pebble) = 1), except for the wood stump (P_f(wood stump) = .53). The other class (14% of the children, Strategy 3b) shows chance level scores for all objects (P_r(boat) = .53, P_r(wood stump) = .40, P_r(coin) = .42, P_r(ball) = .48, P_r(pebble) = .47). Responses of children in this class are around chance level. Class membership is related to age (Table 2.3, Figure 2.2c) as shown by logistic regression (Cox and Snell R^2 = .130; β_{age} = .533, p = .001) relatively to the residual strategy.

Relations between strategies for different sets of objects. To test whether children are consistent in strategy use between different sets we applied a binomial stepwise logistic regression of strategy use for sets 2 and 3 with age and strategies of Set 1 as predictor variables (age entered first). The contribution of Set 1 strategy on top of age is significant in predicting Set 2 strategy (Log Likelihood ratio: $\chi 2$ (2) = 18.299, p < .001), see Table 2.4. With age the probability of the most advanced Set 2 strategy (Strategy 2a) increases. Moreover, children with the two most advanced Set 1 strategies (Strategy 1a and Strategy 1b) have an increased probability of applying Strategy 2a for Set 2 (relative to Strategy 2b). Strategy use for predicting buoyancy of Set 3 objects significantly depends on age (Log Likelihood ratio: $\chi 2$ (1) = 19.357, p < .001). The contribution of Set 1 strategy on top of age is not significant in predicting Set 3 strategies.

Table 2.4
B values of binomial stepwise logistic regression model of strategy use for Set 2.

95% CI for Odds Ratio					
	B (*SE*)	Lower	Upper	p	Exp (*B*)
Constant	1.508	-2.492	7.046	.487	4.520
Age*	.400	.129	.797	.005	1.491
Strategy Set 1	-1.379	-2.999	- .464	.003	0.252

Conclusion prediction phase. To predict buoyancy for indistinctive cubes (Set 1), material cubes (Set 2), and exemplar objects (Set 3) most children used a systematic strategy. Cubes with an indistinctive appearance (Set 1) can only be judged by taking mass and volume in consideration, and therefore were most difficult to judge. Two latent groups of children integrated both dimensions to make their predictions, but differed in their bias towards underestimating mass for small cubes (with density near 1.0) and overestimating mass for large cubes (with density near 1.0). A third group of children did not make systematic predictions at all. In general, older children used more advanced strategies resulting in better performance. Cubes made from the distinctive materials, wood and metal (Set 2) were easier to judge. Two latent groups of children were distinguished, the largest group was quite accurate but appeared to integrate both the mass and the volume of cubes, instead of only looking at material to make predictions. This was also due to the fact that they were underestimating mass for small cubes and overestimating mass for large cubes. Again a smaller group of children (between 4 and 7 years old) did not make systematic predictions. The buoyancy of Exemplar objects (Set 3) was most easy to judge, except for the wood stump. For Set 3 too, a minority of the children did not make systematic predictions. Strategy use for sets 1 and 2, but not sets 1 and 3, was related after correcting for age.

2.3.2 Explanation Phase

Comparing explanations for different sets. First, to compare explanations between sets, children's explanations were scored into 5 categories on a scale (scientific, mass and volume, volume, mass, residual; see Methods). The inter-rater reliability was found to be sufficient (Sattler, 2002): Kappa = 0.91 ($p < 0.001$), 95% CI (0.857, 0.959). Per set (Set 1, 2 and 3) each child was assigned to one, the most advanced type of explanation (Table 2.5). Mentioning material was not scored separately here, because the position on the scale would be unclear for sets 1 and 3. Most children (62%) gave an explanation based on only one dimension (volume or mass) for at least one set. Only 22% of the children gave an explanation for buoyancy based both on mass and volume (correct or incorrect integration).

Table 2.5
Percentages (numbers) of children using types of explanations for the three sets of objects.

Set	Type of Explanation					
	scientific	mass+vol	vol	mass	residual	sum
Indistinctive cubes (Set 1)	10 (14)	17 (23)	2 (3)	61 (85)	10 (14)	100 (139)
Material cubes (Set 2)	6 (8)	14 (19)	6 (8)	58 (80)	17 (24)	100 (139)
Exemplar objects (Set 3)	14 (20)	7 (10)	8 (11)	53 (73)	18 (25)	100 (139)
Sum	10 (42)	12 (52)	5 (22)	57 (238)	15 (63)	100 (417)

Note. scientific indicates an explanation mentioning a scientifically correct combination of mass and volume, mass+vol indicates mentioning both mass and volume, vol indicates mentioning only volume, mass indicates mentioning only mass, residual indicates that the explanation did not match any of the former categories.

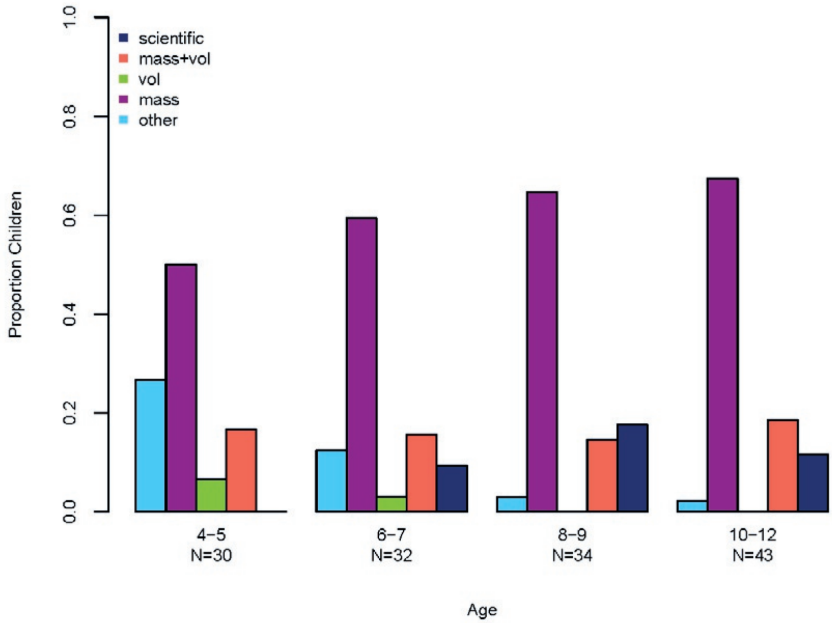


Figure 2.4. Proportion of children explaining buoyancy of Indistinctive cubes (Set 1) within each age group.

To test whether age is related to the type of explanation that children gave, we used age as predictor in a multinomial logistic regression. Age is a significant predictor for the type of explanation given for Set 1, $\chi2$ (3) = 15.875, p = .001, but not for Set 2 and Set 3. To test whether explanations for different sets were related, we used Set 1 explanations as a predictor variable in a multinomial logistic regression. The type of explanations given for buoyancy of Set 1 cubes predicted the type of explanation given for Set 2, $\chi2$ (3) = 14.786, $p < .005$. Also explanations for buoyancy of Set 3 objects were significantly predicted by explanations given for Set 1, $\chi2$ (3) = 9.520, $p < .05$.

Comparing predictions and explanations. Is there a relation between children's strategy use for predicting buoyancy and children's explanations*?* To do this analysis we added a most advanced type of explanation (i.e., the typical explanation) for Set 2: mentioning material (wood and metal). For the scoring of Set 3 explanations, we added 'relevant facts' (i.e., the typical explanation) as the most advanced strategy. Each child has only one explanation category per set. As children often had multiple explanations for one set of objects, the most advanced explanation was assigned (in Table 2.5 from left to right). None of the children used relevant facts in the context of Set 1[5].

[5] Four children mentioned material in the context of Set 1, but were categorized as scientific (1), mass (2), mass+vol (1). 22 children mentioned specific facts as one of the explanations for buoyancy of Set 2 objects, but were categorized as scientific (1), mass+vol (1), vol (2), mass (13) and residual (5). 57 children named 'material' as one of the explanations for buoyancy of Set 3 objects, but were categorized as scientific (10), mass+vol (6), mass (29) and residual (10).

Table 2.6 shows the cross-tables of prediction strategies and explanations. To analyze the relation between prediction strategy and type of explanation, for each set separately we used age and strategy use as predictors in a multinomial stepwise logistic regression to predict the type of explanation given. The contribution of age is significant for predicting explanations of indistinctive cubes of Set 1 (Log Likelihood ratio: $\chi2$ (3) = 15.875, p = .001), older children used more sophisticated explanations than younger children. For Set 1, the contribution of children's prediction strategy on top of age is not significant in predicting explanations. Explanations of buoyancy of Set 2's wooden and metal cubes are only marginally related to age, Log Likelihood ratio: $\chi2$ (3) = 6.333, p = .097. For Set 3, both age and prediction strategy were unrelated to explanations given.

Conclusion explanation phase. The most common explanation that children gave for the buoyancy of cubes with an indistinctive appearance (Set 1) was based only on mass. Mentioning both mass and volume did occur but mostly without integrating both dimensions. In the context of wooden and metal cubes (Set 2) material is mentioned frequently. In the context of exemplar floaters and sinkers (Set 3) children frequently mentioned specific facts. In general, older children gave more advanced explanations, but large individual differences were observed within age groups as well. There is no clear relationship between children's predictions and explanations beyond age.

Table 2.6
Percentages of children using types of explanations for each prediction strategy per set of objects.

Strategy	Type of Explanation							
	fact	material	scientific	mass+vol	vol	mass	residual	sum
Strategy 1a			11 (5)	26 (12)	2 (1)	62 (29)	0 (0)	100 (47)
Strategy 1b			15 (7)	13 (6)	0 (0)	65 (30)	7 (3)	100 (46)
Strategy 1c			4 (2)	11 (5)	4 (2)	57 (26)	24 (11)	100 (46)
Sum Set 1			10 (14)	17 (23)	2 (3)	61 (85)	10 (14)	100 (139)
Strategy 2a		76 (81)	1 (1)	6 (6)	1 (1)	14 (15)	2 (2)	100 (106)
Strategy 2b		21 (7)	0 (0)	6 (2)	9 (3)	45 (15)	18 (6)	100 (33)
Sum Set 2		63 (88)	1 (1)	6 (8)	3 (4)	22 (30)	6 (8)	100 (139)
Strategy 3a	68 (82)		6 (7)	3 (4)	3 (3)	17 (20)	3 (4)	100 (120)
Strategy 3b	53 (10)		0 (0)	5 (1)	11 (2)	21 (4)	11 (2)	100 (19)
Sum Set 3	66 (92)		5 (7)	4 (5)	4 (5)	17 (24)	4 (6)	100 (139)

Note. The first column contains children's strategies for Set 1 (upper part), Set 2 (middle part), and Set 3 (lower part). In columns 2 to 8 seven types of explanations are listed (see text). Each cell contains the percentage (and number) of children that gave a particular explanation.

2.4 DISCUSSION

The aim of this paper was to reveal how children represented knowledge about buoyancy and the conditions under which different types of representations might be triggered. The different conditions tested are different types of objects appearing in a prediction task and an explanation task: objects that have not (indistinct cubes, Set 1) or have some specific association with floating or sinking due to their material appearance (wooden and metal cubes, Set 2) or their object appearance (exemplar objects, such as, a boat and a pebble, Set 3).

2.4.1 Representations of Buoyancy (RQ-1)

Consistent with our first hypothesis, for most children, we found for Set 1 objects information-integration representations. That is, when *predicting* buoyancy for Set 1 objects, most children used a two-dimensional strategy integrating both mass and volume (67% distributed over two systematic strategies versus 33% who made no systematic predictions). The latent class regression model shows that the two systematic strategies were both biased in underestimating mass for small sinkers and overestimating mass for large floaters (with density near 1.0). One strategy was more biased than the other strategy. When *explaining* buoyancy for Set 1 objects on the other hand, 63% of children used a one-dimensional strategy. Hence, as expected, the results show inconsistency in solution strategies between predicting and explaining, which indicates an information-integration-based instead of a rule-based knowledge representation of conceptual knowledge applied in predicting buoyancy of indistinctive cubes.

The analysis of solution strategies for wooden and metal cubes (Set 2) lead to similar conclusions. Part large part of all children (76%, N = 106) used a multi-dimensional solution strategy for the *prediction* task, integrating mass and volume (and possibly in addition also material; the detection of a separate material strategy might depend on the power of the statistical analysis, Van der Maas & Straatemeier, 2008). The other detected solution strategy (24% of the children) was not related to any cube characteristic. Surprisingly, 76% (N = 81) of the children following a multi-dimensional solution strategy for *predicting* buoyancy of Set 2 objects, based their *explanations* on cube material, that is a perfect one-dimensional rule. Hence, a fully accurate rule (cubes made of wood float, cubes made of metal sink) was apparently known to many children, but was not relied on completely while making predictions.

Finally, for Set 3, the largest part of the children (86%) made highly accurate *predictions*, also for the small sinkers (the pebble and the coin, .99 and 1 respectively). 82% of these children mentioned relevant facts during the *explanation* task, which explains the high accuracy of the predictions.

We conclude that for *predicting* buoyancy of cubes, children were triggered to use multi-dimensional information-integration strategies. On the other hand, many children (Set 1 63%, Set 2 88%) based *explanations* on mass only or cube material, which could be considered a rule-based representation. Following Dienes and Perner (1999), one can conclude that a considerable group of children (43%, Table 2.6) shows evidence for having both knowledge that remains implicit in functional use and knowledge that is explicit, such that it is verbalizable. This conclusion is supported by findings in the literature concerning science learning (Butts et al., 1993; Hsin & Wu, 2011; Smith et al., 1985; Smith et al., 1997; Tenenbaum et al., 2004).

The evidence for the information-integrating representations of children's conceptual knowledge of a physical principle adds to the literature in two ways. First, the application of latent regression models to the prediction data made it possible to pitch the information integrating representation against the rule-based representation for the predictions, while at the same time accounting for individual differences in solution strategies. This is unique for the domain of buoyancy (see Hofman et al., 2015, for the domain of torque), and avoids the matching of observed response patterns with expected response patterns, which is methodologically problematic (Van der Maas & Straatemeier, 2008). The latent variable modeling clearly showed that the information-integration representation was superior in describing the prediction data. In addition, latent variable modeling made it possible to isolate responses of children who responded around chance level. That is, among the children up to 7 years old a considerable group (33%, 24%, and 14% for the successive sets) did not make any systematic predictions. Analysis in future studies should account for this group of children as well.

Second, the current study used 3D objects that were handed over to the children, instead of descriptions or drawings of objects (Hardy et al., 2006). 3D objects relate much more directly to children's informal experiences. Hence, one would expect a 3D task to be a more valid measure of conceptual knowledge as it is developed in everyday experiences. A disadvantage of testing with 3D objects might be that children (and adults) integrate weight and volume automatically when perceiving weight and volume directly. It has been shown that in estimating the mass of an object its volume might interfere, which is known as the size-weight illusion (Dijker, 2014). Strikingly, the overestimation of the importance of weight, which children showed in the *prediction* task, appeared to be in the opposite direction of the size-weight illusion. That is, the latent regression model shows that children were not overestimating mass for small objects, but were instead biased to predict floating for smaller, sinking cubes. Moreover, children were not underestimating mass for large objects, but were biased to predict sinking for larger floating cubes. Hence, the integration of information for predicting buoyancy could not be explained by the size-weight illusion.

The revealed differences in knowledge representations between tasks and between types of objects are potentially very interesting for an educational context. The current study suggests that in order to trigger children's knowledge representation based on the integration of mass and volume, the presentation of cubes with different density, volume and mass, and with equal, indistinctive appearance would be most effective. A buoyancy lesson with such objects would result in an inconsistency between children's predictions and explanations, which is an interesting starting point for learning (Carey, 2000; Van Schijndel, Visser, Van Bers, & Raijmakers, 2015).

2.4.2 Development of Representations (RQ-2)

How do these solution strategies develop with age? First of all, results of the prediction task show that the integration of information improved with age (see also Janke, 1995; Wilkening & Anderson, 1982). That is, regardless that most of the older children are still *explaining* buoyancy by a one-dimensional rule (mass or material), they integrated information more accurately than younger children while making *predictions*.

Results of the explanation task show developmental change less clearly. Indeed, the type of explanation for the buoyancy of indistinctive cubes changed with age, but this was largely due to a decrease of the frequency of ‘other’ (i.e., irrelevant) explanations. Surprisingly, the number of one-dimensional explanations was generally high and did not decrease with age. Hence, we did not find evidence that developmental changes in explanations of buoyancy were developing from one-dimensional to two-dimensional, compound rules (Siegler et al., 1981) or scientific concepts (Schneider & Hardy, 2013). A reasonable explanation would be that advanced insight in buoyancy only arises as a result of formal education, as was suggested by Leuchter et al. (2014), and children in the Netherlands typically receive formal instruction about buoyancy only in secondary school.

Explanations for the exemplar objects also do not show a clear, developmental pattern. For the Exemplar objects, many children explained buoyancy by referring to specific relevant facts, such as “pebbles are on the bottom of a river” (82% of the children who made high-accuracy predictions). The use of specific facts might agree best with Schneider and Hardy’s (2013) everyday profile of conceptual knowledge of buoyancy, because it can coherently explain a set of observations from everyday life although the possibility to generalize the explanation is limited.

2.4.3 Limitations of this Study

In this study we compared results between different tasks, the prediction and the explanation task, and between sets of objects, indistinctive cubes, material cubes, and exemplar objects. The way we organized these conditions within participants might be reason to be cautious about some results. First, in the prediction task each object was presented separately. In contrast, during the explanation task for each set two baskets were presented filled with the objects sorted by the child as floating and sinking. The reason for not asking explanations for each object separately after predicting its buoyancy, is that self-explanations might affect representations and therefore future predictions (Williams & Lombrozo, 2010). Moreover, in the explanations task we intended to optimize the situation for children to mention mass and volume in combination to each other by showing them that they categorized objects with different characteristics together. Anyway, it appeared that despite this combined presentation, 56% of the children did discuss multiple objects separately during the explanation task of Set 1. Hence, we would expect that an explanation based on mass only would occur even more frequently in a different set up. It would be interesting to test in future research whether and how explanations depend on the variation of objects shown at the same time.

Second, the order of presentation of the different sets was fixed, such that the number of relevant object characteristics increased during the experiment. In this manner, we aimed at triggering relevant object characteristics by the objects and not by the context (i.e., the earlier presented objects). It would be interesting to test the stability of children's explanations and predictions by systematically manipulating the presentation sequence.

2.5 CONCLUSION

Analyzing predictions and explanations of children between 4 and 12 years old, separately for different types of objects revealed new insights in children's knowledge representations about buoyancy of objects. As was expected from the perspective of categorization learning research, multiple knowledge representations within and between individuals were found, depending on task demands and types of objects to classify. In predicting buoyancy, children integrated mass and volume when they did not have other object characteristics to rely on. Even when children gave a correct explanation based on material, they relied on integrating mass and volume when predicting buoyancy. For exemplar objects, children explained buoyancy largely by specific facts. Clear development was shown for prediction strategies (but not for the explanations), such that the way mass and volume were integrated improved with age (see also, Wilkening & Anderson, 1982). Finally, a group of children (only present among 4- to 7-year-olds) was not systematic in predicting buoyancy of objects, which should be accounted for in future studies.

2.6 SUPPLEMENTARY MATERIAL

2.6.1 Results

Children's categorical responses (float, sink) of the prediction phase were modeled with latent class regression models (Set 1 and Set 2) and latent class models (Set 3). To find the best fitting, most parsimonious model exploratory, base models with increasing complexity were fit at the data of each set separately (see methods section on analysis in the main article). Below, we present the fit statistics of the models that were considered in the model selection process (Set 1: Supplementary Table S2.1; Set 2: Supplementary Table S2.2; Set 3: Supplementary Table S2.3).

Table S2.1

Fit statistics for Latent Class Regression Models for Set 1.

Model	#class	LogL	#par	AIC	BIC
1c- $\beta_{mass} + \beta_{vol}$	1	-607.72	3	1221.44	1236.45
2c- $\beta_{mass} + \beta_{vol}$	2	-536.74	5	1083.48	1108.49
3c- $\beta_{mass} + \beta_{vol}$	3	-518.46	11	1058.91	1113.95
2c- $\beta_{mass} + \beta_{vol}$ & 1c- a	3	-520.64	9	1059.28	1104.31
2c- $\beta_{mass} + \beta_{vol}$ & 1c- β_{mass}	3	-518.87	10	1057.74	1107.77
2c- $\beta_{mass} + \beta_{vol}$ & 1c- β_{vol}	3	-520.1	10	1060.21	1110.24
2c- $\beta_{mass} + \beta_{vol}$ & 1c- β_{color}*	3	-504.28	12	1032.57	1092.60

Note. The models describe children's prediction responses of the unfamiliar cubes (Set 1). LogL indicates log likelihood; #par, number of estimated parameters; AIC, Aikaike's Information Criterion; BIC, Bayesian Information Criterion. Model 1c- $\beta_{mass} + \beta_{vol}$ (2c- $\beta_{mass} + \beta_{vol}$, 3c- $\beta_{mass} + \beta_{vol}$) indicates a single- (two-, three-) class base model with effects for log(mass) and log(edge length), in addition to an intercept. Model 2c- $\beta_{mass} + \beta_{vol}$ & 1c- a (or & β_{mass}, & β_{vol}, & β_{color}) consists of two classes with effects for log(mass) and log(edge length) and one class with only an intercept (or an intercept + effects for log(mass), log(edge length), cube color, respectively). Comparable 2-class models were less parsimonious than the presented 3-class models. An asterisk (*) indicates the most parsimonious, best-fitting model.

Table S2.2

Fit statistics for Latent Class Regression Models for Set 2.

Model	#class	LogL	#par	AIC	BIC
1c- $\beta_{mass} + \beta_{vol}$	1	-360.23	3	726.45	740.62
2c- $\beta_{mass} + \beta_{vol}$	2	-322.45	7	658.89	691.95
3c- $\beta_{mass} + \beta_{vol}$	3	-316.63	11	655.26	707.21
1c- $\beta_{mass} + \beta_{vol}$ & 1c- a*	2	-324.46	3	654.92	669.09
1c- $\beta_{mass} + \beta_{vol}$ & 1c- β_{mass}	2	-324.1	6	660.2	688.53
1c- $\beta_{mass} + \beta_{vol}$ & 1c- β_{vol}	2	-323.31	4	654.61	673.5
1c- $\beta_{mass} + \beta_{vol}$ & 1c- β_{color}	2	-323.72	6	659.43	687.77
1c- β_{color} & 1c- a	2	-565.22	2	1134.42	1143.89
1c- $\beta_{mass} + \beta_{vol} + \beta_{color}$ & 1c- a	2	-321.70	4	651.41	670.30

Note. Fit statistics of the models describing children's prediction responses of the wooden and metal cubes (Set 2). See Supplementary Table S2.1 for explanation of the symbols. Model 1c- $\beta_{mass} + \beta_{vol}$ (2c- $\beta_{mass} + \beta_{vol}$, 3c- $\beta_{mass} + \beta_{vol}$) indicates a single-(two-, three-) class base model with effects for log(mass) and log(edge length), in addition to an intercept. Model 2c- $\beta_{mass} + \beta_{vol}$ & 1c- a (or & 1c - β_{mass}, & 1c- β_{vol}, & 1c- β_{color}) consists of two classes with effects for log(mass) and log(edge length) and one class with only an intercept (or an intercept + effects for log(mass), log(edge length), cube material, respectively). Model 1c- β_{color} & 1c- a consists of one class with an effect for cube material and one class with only an intercept. Model 1c- $\beta_{mass} + \beta_{vol} + \beta_{color}$ & 1c- a consists of one class with an effect for mass, volume, and cube material and one class with only an intercept. An asterisk (*) indicates the most parsimonious, best-fitting model.

Table S2.3
Fit statistics for Latent Class Models for Set 3.

Model	LogL	df	AIC	BIC
1 class	-275.13	5	560.27	574.94
2 class*	-246.47	11	514.94	547.22
3 class	-243.99	17	521.98	571.86
1 class-age	-257.06	10	534.11	563.46

Note. The models describe children's prediction responses of the familiar objects in Set 3. See Supplementary Table S2.1 for explanation of the symbols. Model 1 class-age is a one-class latent class model in which the accuracy of children's responses for each object is regressed on age. An asterisk (*) indicates the most parsimonious, best-fitting model.