British Journal of Educational Psychology (2019)

The British Psychological Society

www.wileyonlinelibrary.com

Children's reasoning about continuous causal processes: The role of verbal and non-verbal ability

Selma Dündar-Coecke[1]*, Andrew Tolmie[1] and Anne Schlottmann[2]

[1]Centre for Educational Neuroscience, Department of Psychology and Human Development, UCL Institute of Education, University College London, UK
[2]Department of Experimental Psychology, University College London, UK

Background. Causes produce effects via underlying mechanisms that must be inferred from observable and unobservable structures. Preschoolers show sensitivity to mechanisms in machine-like systems with perceptually distinct causes and effects, but little is known about how children extend causal reasoning to the natural continuous processes studied in elementary school science, or how other abilities impact on this.

Aims. We investigated the development of children's ability to predict, observe, and explain three causal processes, relevant to physics, biology, and chemistry, taking into account their verbal and non-verbal ability.

Sample. Children aged 5–11 years (N = 107) from London and Oxford, with wide ethnic/linguistic variation, drawn from the middle/upper socioeconomic status (SES) range.

Methods. Children were tested individually on causal tasks focused on sinking, absorption, and dissolving, using a novel approach in which they observed contrasting instances of each, to promote attention to mechanism. Further tasks assessed verbal (expressive vocabulary) and non-verbal (block design) ability.

Results. Reports improved with age, though with differences between tasks. Even young participants gave good descriptions of what they observed. Causal explanations were more strongly related to observation than to prediction from prior knowledge, but developed more slowly. Non-verbal but not generic verbal ability predicted performance.

Conclusions. Reasoning about continuous processes is within the capacity of children from school entry, even using verbal reports, though they find it easier to address more rapid processes. Mechanism inference is uncommon, with non-verbal ability an important influence on progress. Our research is the first to highlight this key factor in children's progress towards thinking about scientific phenomena.

Causal cognition – the ability to perceive and infer cause–effect relations – lies at the core of scientific investigation and is equally crucial in everyday thinking. It revolves around the notion that causes produce effects by means of an underlying mechanism. While some

**Correspondence should be addressed to Selma Dündar-Coecke, Centre for Educational Neuroscience, Department of Psychology and Human Development, UCL Institute of Education, University College London, 25 Woburn Square, London, UK (email: selma.coecke@gmail.com).*

DOI:10.1111/bjep.12287

aspects of a phenomenon are observable, causal cognition goes beyond immediate observation and requires inference from observables to invisible factors.

Although humans employ a variety of sources, such as statistical relations, prior knowledge or temporal information, cuing them to access causal mechanism, two modes of human inference have typically been seen as connecting these observable and unobservable determinants: observation of regularities, and intervention (Danks, 2005; Lagnado & Sloman, 2004; Lagnado, Waldmann, Hagmayer, & Sloman, 2007; Gopnik *et al.*, 2004; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003; Waldmann & Hagmayer, 2005; see also Rescorla & Wagner, 1972; Shanks & Dickinson, 1987; Sloman, 1996, for associative accounts; and e.g., Gopnik *et al.*, 2004, on probabilistic inference; Sloman & Lagnado, 2005 on how counterfactual reasoning is seen as a type of imagined intervention). The question here is whether children can conceive causal mechanisms from observation.

Preschoolers engage in inferences about mechanism in clearly structured causal systems, typically simple machines, provided they have prior knowledge or experience of these (Buchanan & Sobel, 2011; Bullock, Gelman, & Baillargeon, 1982; Schlottmann, 1999; Shultz, 1982). Buchanan and Sobel (2011) showed children that pressing one of two buttons made a light go on; the causal button had a sticker and was connected to the light by a wire or had a battery inside. Four-year-olds predicted that if the wire was switched, the other button would make the light come on; even 3-year-olds predicted this if the battery was switched. Neither age thought switching stickers would affect the outcome, indicating prior knowledge of what was relevant to the mechanism, and concern with how this led from cause to effect beyond what was observed.

However, although such demonstrations show preschool children are able to engage in forms of mechanism-based reasoning, the tasks typically consider the effects of mechanism-related variables on outcome. Children were not asked about, nor understood, presumably, the processes by which these variables exerted their effects, for instance, the operation of electrical circuits or chemical reactions inside the battery. Little is known about how this level of understanding develops, or how children go on to reason causally about less clearly structured or never fully observed processes in the natural world, although they encounter these in much elementary school science. There are related literatures, such as work on children's intuitions regarding physical phenomena (see Wilkening & Cacchione, 2011), but the tasks used in this are designed to elicit pre-existing and implicit knowledge of causal connections, not explicit *reasoning* about processes to further such knowledge.

This study aims to provide the first systematic investigation of such thinking, and whether children can acquire causal knowledge from observation of natural phenomena relying on continuous processes without intervention. We examined the developmental trajectory of 5- to 11-year-olds' predictive, observational, and explanatory competences with regard to reports of three phenomena, relevant to physics, biology, and chemistry (sinking, absorption, and solution), chosen to encompass a manageable range of familiar everyday phenomena. Using a new paradigm, children were asked to consider directly contrasting instances of each phenomenon, to draw their attention to the mechanisms involved (See Dündar-Coecke & Tolmie, 2019a, in press) on the benefit of contrasting instances in revealing causal relationships). Our objective was to determine the extent to which they were able to engage with thinking about causal mechanisms, and how far this related to their prior knowledge and current observation of the phenomena, and to their verbal and non-verbal ability.

From causal events to causal processes

Causal reasoning tasks in both the developmental and adult literature typically involve simple machines, that is physical (or virtual) apparatus with distinct components yielding a well-defined, segmented sequence of events. Such structure helps operationalize what may be the cause, mediating mechanism and effect, allowing experimental manipulation of what can be observed and has to be inferred. Other adult studies involve descriptions of causal sequences (Buchsbaum, Griffiths, Plunkett, Gopnik, & Baldwin, 2015; Talmy, 1988; White, 2014), which help structure events that in real life may be more continuous and less easy to segment. Some paradigms in the tradition of Michotte (1946/1963; Schlottmann, Ray, Mitchell, & Demetriou, 2006) involve reasoning about brief events, mostly collisions, which also provide natural segmentation. This schema of causality between distinct, clearly segmented events is common to the vast majority of psychological studies.

However, some forms of causality seem excluded by this. In particular, natural systems typically involve temporally more extended continuous processes, as when an object sinks, dissolves, or absorbs water. Here, there are no distinct cause and effect events; a standard causality-as-events approach would imply that the hand letting go of the stone is the event that causes its sinking, but this is intuitively 'not the right cause' and would seem better regarded as an enabling condition. Salmon (1984) distinguished causality in such continuous processes from causal interactions, though the distinction is not absolute (Kitcher, 1989); for example, sinking involves an interaction of a stone subject to gravity with water pushing upwards. However, while processes can – and perhaps need to – be considered as chains of interactions, no interaction is actually perceived – one *sees* the stone sink, but must *think* about the role of the water. Observers therefore do not see the sinking as either caused by a distinct event via a mediating mechanism, or as an interaction between stone and water; the phenomenal experience presented by the causal sequence is of a uniform, unsegmented process caused by a continuously operating underlying mechanism – the stone sinks because the buoyancy of the water holding it up is less than the weight of the object pulling it down.

Reasoning about causality-as-events, interactions and causality-as-processes all aim to distinguish true from spurious or pseudo-causality by appeal to an underlying mechanism. In each case, children must reason beyond what may be observed at the perceptual surface. However, the distinct components of machines typically reveal the operation of the underlying mechanism more clearly, while it may require more effort to go beyond the perceptual surface of continuous processes. We argue therefore that reasoning about causal mechanisms underlying the simple physical processes encountered in elementary science is a step up from causal reasoning about the toy-machine systems previously studied, and investigate here children's ability to extend their causal thinking to these.

In doing so, we focus on children's capacity to address variation in a single feature of each process, its speed. We focus on this because speed is a characteristic feature of continuous processes. Processes differ in the speed with which they unfold – sinking is generally faster than absorption, which in turn is faster than solution – and this may affect children's ability to observe and infer from observation. More importantly, though, there are differences in speed within each process, depending on the objects involved – a stone sinks faster than a berry, for instance. The cause of this difference relates to the underlying mechanism; for example, the stone sinks faster because the imbalance between liquid and object density is greater. Witnessing this variation in speed alters the experience from a simple instance of a class of phenomena to one of related but quantitatively contrasting instances, leaving this difference to be explained. The use of minimally contrasting cases to highlight a problem dimension derives from perceptual psychology (Gibson & Gibson, 1955) and is frequent in science education (Schwartz, Chase, Oppezzo, & Chin, 2011).

Here, these contrasts should help children structure their observations, orienting them towards the target mechanism in an otherwise uniform process.

Causal reasoning with words

We focus here on how children reason about observed causal processes in their verbal reports. Looking at their causal reasoning in this way is novel, as we moved from the largely non-verbal responses used in preschool causal studies to explanations, to provide insight into how causal reasoning develops towards more demanding forms.

Asking children for explanations about causal phenomena through direct questions has been common since Piaget (1929; see his open-ended questions about natural phenomena, e.g., where the sun came from; and similarly Gelman & Kremer, 1991, tasks combining 'what' and 'why' questions distinguishing natural phenomena from human-made artefacts). This work documents that children can give reasonable explanations of simple events from the preschool age (e.g., why leaves turn brown), and that explanations increase with age in sophistication and accuracy (Vosniadou, 2014; Zaitchik, Solomon, Tardiff, & Bascandziev, 2016). Other studies show that children are active causal explanation-seekers (Carey, 1985; Gopnik & Meltzoff, 1997; Legare, 2014; Legare & Lombrozo, 2014; Walker, Lombrozo, Williams, Rafferty, & Gopnik, 2017; Wellman & Gelman, 1998).

The difference between previously employed methods and those used here is that real phenomena were demonstrated live using direct comparison, allowing children to observe the causal processes. Their explanations did not rely purely on prior knowledge; therefore, the aim was to ask them to make inferences from their observations, following the approach generally adopted in causal reasoning research. We elicited verbal reports of these demonstrations in structured fashion, separating explanation, as a measure of ability to reason beyond what was observed, from prediction based on prior knowledge and description of current observation. We also followed up their initial answers, to minimize verbal demands and response ambiguities by posing further structured questions and by supporting children's expression of knowledge through careful probing.

The measure of explanation was designed to distinguish between children who could only identify simple factors at work in the observed phenomena (e.g., weight), those who recognized that variation in these factors is associated with differences in speed of process (i.e., who treat them as variables), and those who were able to move beyond variables *per se*, to think about the underlying mechanisms connecting variables to outcomes. The measures of prior knowledge and description made it possible to assess how far level of explanation was a function of existing knowledge versus current observation.

Nevertheless, we were aware that apparent quality of explanation might reflect limited language not limited reasoning: although children's understanding of natural causal events is often sophisticated, they may perform poorly compared to adults at expressing and assessing their own causal knowledge (Legare & Clegg, 2014). It is for this reason that developmental psychologists in recent years have attempted to capture their understanding using non-verbal tasks. To address this, measures of language and non-verbal ability were included, as a direct test of how far children's explanations are specifically a function of verbal ability.

Method

Participants

We recruited 120 children, with parental consent, from three primary schools in London and Oxford. We subsequently excluded 13 participants for low attention span or unwillingness to continue; one child was excluded following parental request, leaving 107 individuals for analysis (35 from Year 1 [Y1], mean age = 6 years, 1 month, *SD* = 4.4 months; 33 from Year 3 [Y3], mean age = 8 years, 4 months, *SD* = 5.9 months; 39 from Year 5 [Y5], mean age = 10 years, 3 months, *SD* = 5.9 months).

Of 103 participants whose parents responded, 59 (55.1%) were from monolingual (English) environments, and 44 (41.1%) from bilingual/trilingual homes, confirming the sample encompassed wide ethnic/linguistic variation. The sample was skewed towards the upper ranges of socioeconomic status (SES), with 8 (7.5%) parents being manual workers, 17 (15.9%) self-employed/non-manual workers, and 78 (72.9%) professionals. One parent (0.9%) had only GCSE (basic secondary school) qualifications; 6 (5.6%) had A levels (higher secondary); 39 (36.4%) had undergraduate degrees; 33 (30.8%) had postgraduate degrees or professional qualifications; and 24 (23.3%) had doctoral degrees.

Materials and procedure

Tasks were given to children in fixed order, as below, within a single one-to-one session, in a quiet location within school. Sessions lasted on average 37 min (min = 18, max = 45). Responses were recorded manually on score sheets, but children's replies during the causal reasoning tasks were also audio-recorded for later checking.

The *expressive vocabulary* and *block design* subtests from the Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 2011) were used to provide measures of verbal and non-verbal ability. Administration and scoring followed standard procedures.

The three *causal tasks* each highlighted a direct contrast between two instances of the target phenomenon, presented simultaneously. For *sinking*, children saw a stone and a blueberry of similar size but different densities, which sank at different rates in a large jar of water. For *absorption*, they saw water rising from a Petri dish through strips of tissue and blotting paper of the same length/width, the water rising faster through the more open structure of the tissue. For *solution*, children saw the same small quantities of table and rock salt dissolve in warm water, the greater surface area to volume of the table salt leading to more rapid solution (see Figure 1).

Each task had the same three-stage structure, in which children: (1) inspected the contrasting materials and were asked what they thought would happen when they were put into the water (prediction from *prior knowledge*); (2) watched the focal events and were asked to *describe* what they had noticed; and (3) were asked to *explain* why they thought things had happened in the way that they had seen. At each stage, they were encouraged to give as full an answer as they could (e.g., 'do you think the same thing will happen to both?'; 'did you notice anything else?'; 'do you think there might be another reason?').

Responses were given initial scores for prior knowledge, description, and explanation as children answered (see Table 1 for scoring system). Prior knowledge and description were scored for accuracy of anticipating/reporting differences in sinking/absorption/solution rate. Explanation scoring began at the minimal level of the observed factor(s) (score of 1); via making explicit that these are variables linked to the observed differences in speed of the contrasting examples (score of 2); to a statement about the underlying mechanism which produced the effect (score of 3) (Example responses are shown in Appendix).

To confirm reliability, two researchers subsequently scored all responses independently from the audio-recordings. Agreement was 93%, and final scores were assigned following discussion and checking the audios in the small number of instances where there was a difference.

Tasks	Materials
Sinking (stone, blueberry)	
Absorption (blotting paper, tissue paper)	
Solution (table salt, rock salt)	

Figure 1. Materials used in the causal tasks.

Composite scores were computed for each task (0–7), for each response component (0–6 for prior knowledge and description; 0–9 for explanation), and for number of mechanism-level responses across tasks (0–3).

Results

Analyses utilized data from the 107 participants who completed testing appropriately, except where noted. All tests were two-sided where relevant with $p < .05$. Observed power for regression analyses was .95.

Causal task performance

Tasks

Figure 2 shows the response profiles for each age group on the sinking, absorption, and solution tasks, based on composite scores across prior knowledge, description, and explanation. Performance was best on sinking, followed by absorption, with solution some

Table 1. Scoring system for causal tasks

Component	Sinking	Absorption	Solution
Prediction from *prior knowledge* (0–2)	Correct prediction for stone (i.e., sinks) = 1 Correct prediction for difference between stone and berry (i.e., sink at different speeds) = 1	Correct prediction for tissue paper = 1 Correct prediction for difference between tissue and blotting paper = 1	Correct prediction for table salt = 1 Correct prediction for difference between table and rock salt = 1
Description of observation (0–2)	Correct description for stone = 1 Correct description for berry = 1	Correct description for tissue paper = 1 Correct description for blotting paper = 1	Correct description for table salt = 1 Correct description for rock salt = 1
Explanation/inference (0–3)	No/irrelevant explanation = 0 Weight/size without difference between objects = 1 Weight/size with difference = 2 Density and mechanism = 3	No/irrelevant explanation = 0 Thickness/softness/texture etc. without difference between types of paper = 1 Thickness/softness/texture etc. with difference = 2 Nature of papers/holes and mechanism = 3	No/irrelevant explanation = 0 Grain/size etc. without difference between types of salt = 1 Grain/size etc. with difference = 2 Grain/size etc. with surface area and mechanism = 3

way behind. A two-way mixed ANOVA (task within-subjects, age between-subjects) found a significant main effect of task, $F(2, 208) = 47.202$, $p < .001$, partial eta-squared = .312, with no significant difference between scores on sinking and absorption using a Bonferroni comparison, but with both significantly higher than scores for solution. There was also a main effect of age group, $F(2, 104) = 24.250$, $p < .001$, partial eta-squared = .318, with scores for Y1 significantly lower than Y3, but no difference between Y3 and Y5; and a modest task $\times$ age interaction, $F(4, 208) = 5.056$, $p = .001$, partial eta-squared = .089, reflecting greater growth on solution between Y1 and Y3 than for sinking and absorption.

Causal components

The profiles of each age group for prior knowledge, description, and causal explanation across the three causal tasks are shown in Figure 3. Children performed at a high level on description, slightly less well on prior knowledge, and at a notably lower level on causal explanation. For *description*, 91.6% of children obtained the maximum score in the sinking task, 85% in absorption, and 57.9% in solution. For *prior knowledge*, the corresponding values were 72.9% for sinking, 62.6% for absorption, and 37.4% for solution.

In contrast, for *explanation*, only 9.3% got the highest score in sinking, 14% for absorption, and 4.7% for solution. The majority of explanation responses on all three tasks focused solely on identification of causal factors or variables (scores of 1 or 2). Although mechanism responses became more common in the two older groups (for sinking, there were 0 in Y1, 4 in Y3, and 5 in Y5; for absorption, 1, 6, and 8 respectively, and for solution, 0, 0, and 5), children apparently found it difficult to make the shift to this level of thinking.

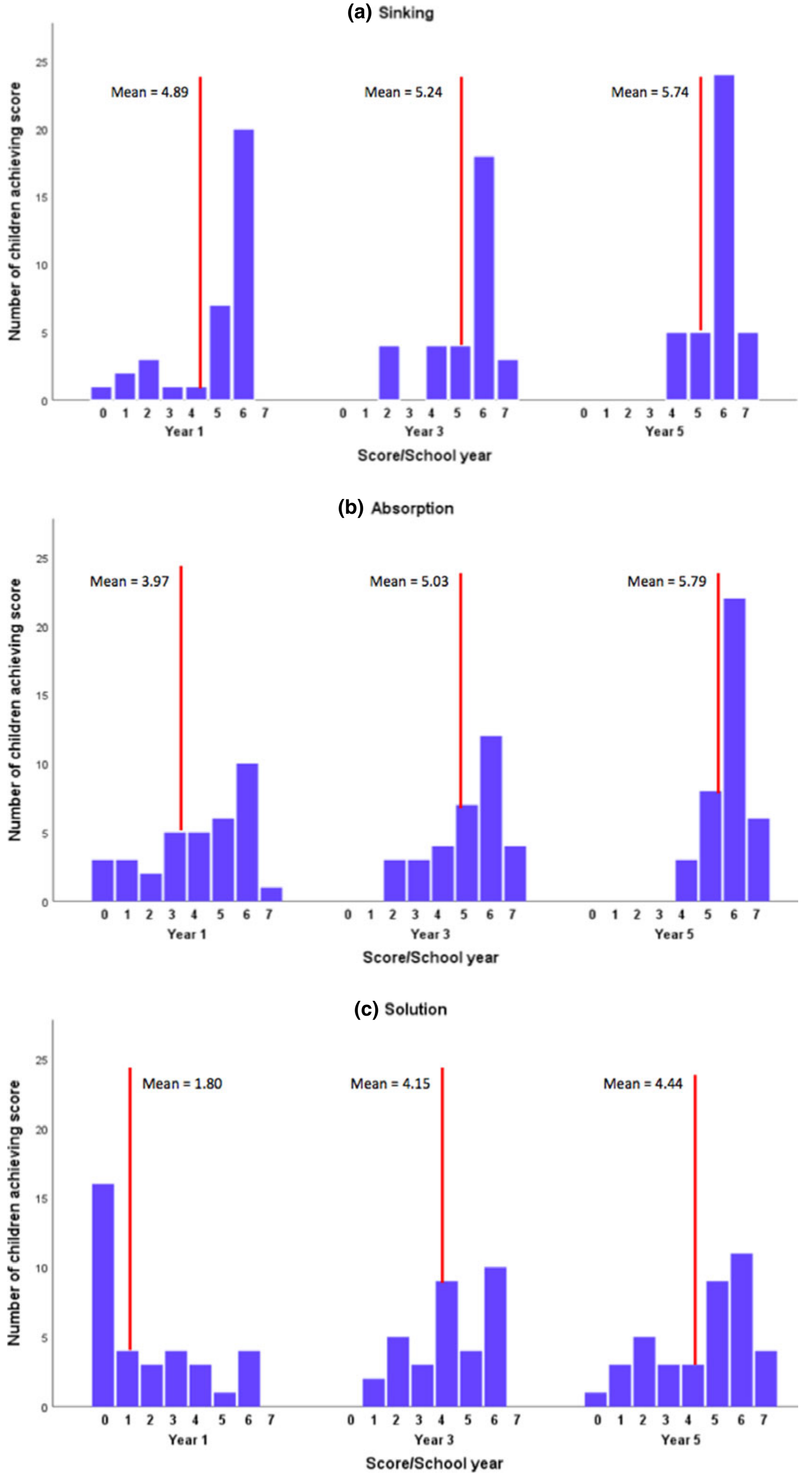

Figure 2. Profile of scores on (a) sinking; (b) absorption; and (c) solution (max = 7).

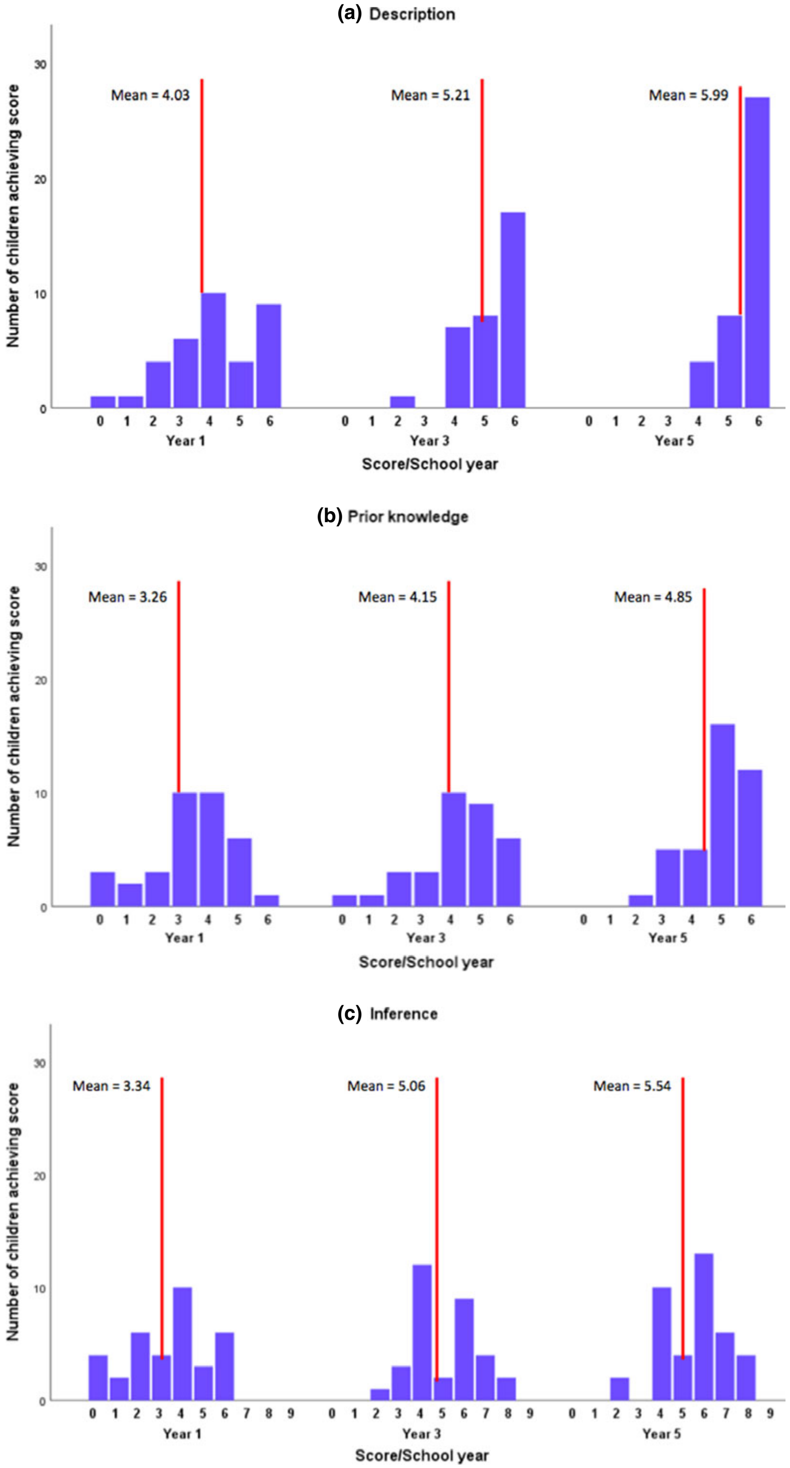

Figure 3. Profile of scores on (a) prior knowledge (max = 6); (b) description (max = 6); and (c) causal explanation (max = 9).

However, if they made any reference to mechanism at all, they tended to do so on more than one task, at over 2.5 times the chance rate, so the shift appeared to be domain-general when it occurred.

Direct statistical comparisons between the components were not made, given differences in the scales and dimensions measured. However, one-way ANOVAs showed age-related progression on each: for *prior knowledge*, $F(2, 104) = 12.376$, $p < .001$, partial eta-squared = .192; for *description*, $F(2, 104) = 18.336$, $p < .001$, partial eta-squared = .261; and for *explanation*, $F(2, 104) = 17.383$, $p < .001$, partial eta-squared = .251. For each component, there were significant differences between Y1 and Y3, but not between Y3 and Y5. The components were positively correlated with each other, controlling for age (for *prior knowledge* and *description*, $r = .392$; for *prior knowledge* and *explanation*, $r = .414$; for *description* and *explanation*, $r = .618$, all $p < .001$).

Is children's causal reasoning associated with verbal and non-verbal ability?

There was significant positive skew on block design, due to the oldest age group having a longer tail. Vocabulary was normally distributed. One-way ANOVAs found significant increases with age on both, however: for vocabulary, Welch robust statistic = 54.093 ($df = 2, 67.790$); for block design, 45.070 (2, 63.948), $p < .001$ for both, with significant differences between all three age groups on both measures: for vocabulary, Y1 mean = 22.89, $SD = 5.290$, Y3 mean = 30.76, $SD = 5.863$, Y5 mean = 35.62, $SD = 5.204$; for block design, Y1 mean = 11.91, $SD = 6.085$, Y3 mean = 19.15, $SD = 9.517$, Y5 mean = 34.19, $SD = 13.250$. Variance was not notably attenuated for either measure: for vocabulary, overall mean = 29.95, $SD = 7.586$; for block design, overall mean = 22.23, $SD = 13.860$.

Correlations between variables

Zero-order Pearson correlations showed the three causal components were positively associated with both vocabulary and block design scores, which were themselves positively correlated with each other (Table 2). The relationship of block design to the causal measures was logarithmic, and log block design (the logarithmic transform) was more strongly correlated with these than the untransformed score. When age in months

Table 2. Zero-order and partial correlations between causal reasoning, verbal, and non-verbal ability (significant associations in bold)

	Prior	Description	Explanation	WASI vocab	Block	Block (log)
Prior	**1**	**.518*****	**.531*****	**.466*****	**.473*****	**.555*****
Description	**.392*****	**1**	**.703*****	**.439*****	**.391*****	**.476*****
Explanation	**.414*****	**.618*****	**1**	**.467*****	**.441*****	**.516*****
WASI vocabulary	**.265****	.150	**.224***	**1**	**.677*****	**.679*****
Block design	**.286****	.104	**.195***	**.416*****	**1**	**.923*****
Block design (logarithmic)	**.408*****	**.241***	**.315*****	**.431*****	**.867*****	**1**

Notes. Zero-order correlations above diagonal, $N = 107$; partial correlations below diagonal, $N = 106$ due to missing date of birth data for one participant.
*$p < .05$; **$p < .01$; ***$p < .001$.

was controlled for, log block design showed a stronger correlation with the components than vocabulary, which was uncorrelated with description responses. Parental occupation and education correlated with each other, $r = .609$, but otherwise only with non-verbal ability, .261, $p = .008$ and .382, $p < .001$, respectively, and are not considered further.

Hierarchical regression models

Hierarchical regressions examined the unique variance accounted for by verbal and non-verbal ability, given the association between them. Taking causal component scores and number of mechanism responses as the dependents, age in months was entered in the first, WASI vocabulary in the second, and log block design at the third stage, to assess whether their effects were distinct from each other.

This analysis produced significant models and final adjusted R-square in all four analyses (Table 3). Vocabulary was a significant predictor at the second stage for prior knowledge and explanation, but not for description, confirming the partial correlation; or for mechanism responses. The inclusion of log block design consistently led to both age and vocabulary dropping out, leaving it the only predictor, with one exception – description, where age remained significant.

Path analysis using a maximum likelihood approach was employed to test the fit of the regression model for mechanism responses, treating age as a background influence. The model in which the influence of vocabulary was entirely mediated by log block design (Figure 4) provided the best fit to the data, chi-square = 0.007, $df = 1$, $p = .932$. A further mediation–moderation analysis confirmed these effects, showing that there was no interaction between vocabulary and blocks ($p > .05$), and that there was full mediation:

Table 3. Hierarchical regression analysis with component causal scores and mechanism responses as dependent variable (significant predictors in bold)

	Model	M1	M2	M3
	Predictor		β	
Prior knowledge	Age in months	**.424*****	.197	.059
	WASI vocabulary		**.330****	.136
	Block design (log)			**.425*****
Adj $R^2 = .304$; $\Delta R^2 = .180$*** for M1; .057** for M2; .087*** for M3				
Description	Age in months	**.502*****	**.379****	**.300***
	WASI vocabulary		.178	.067
	Block design (log)			**.243***
Adj $R^2 = .277$; $\Delta R^2 = .252$*** for M1; .017** for M2; .028* for M3				
Explanation	Age in months	**.472*****	**.285***	.185
	WASI vocabulary		**.272***	.132
	Block design (log)			**.307***
Adj $R^2 = .287$; $\Delta R^2 = .223$*** for M1; .039* for M2; .045* for M3				
Mechanism	Age in months	**.284****	.191	.088
	WASI vocabulary		.135	−.010
	Block design (log)			**.317***
Adj $R^2 = .114$; $\Delta R^2 = .081$** for M1; .010 for M2; .048* for M3.				

Note. *$p < .05$; **$p < .01$; ***$p < .001$.

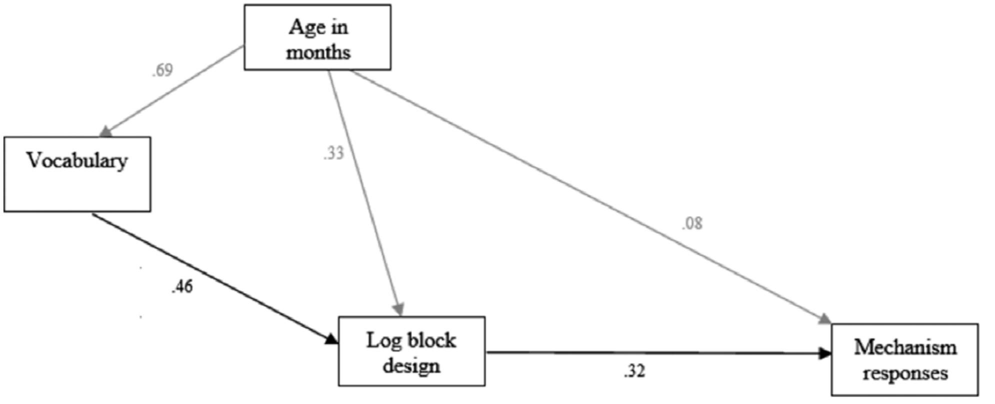

Figure 4. Path model including standardized coefficients for the effects of age, vocabulary, and log block design on mechanism responses (subsidiary relationships in grey).

the direct effect of vocabulary was non-significant, $b = .00$, $SE = .01$, $t = 0.24$, $p = .81$ with a small unstandardized indirect effect $b = .02$, CI = 0.01–0.05, standardized effect = .24, indicating the path model was robust.

Discussion

The development of causal reasoning about continuous processes

Differences between the three causal processes

Children's reports improved with age for all three continuous processes. However, there were task differences as well, with sinking easiest, and solution most difficult. Whether dissolving is more difficult in general for children, or whether the task differences here reflect issues with the specific examples of the processes chosen remains open.

We originally chose these three processes in part for their obvious contrast in relative speed, on the view that faster processes, with lower demands on sustained attention, might be easier for children, and the outcomes were consistent with this. It took 1–3 s for the stone and berry to sink, 3–5 s for the water to rise through the tissue/blotting papers, but 60–120 s for the smaller and larger grains of salt to visibly begin to dissolve. However, there are other possible accounts. Solution was also harder to see because of the smaller scale at which it occurred; it involves a less accessible, sub-microscopic mechanism (Liu & Lesniak, 2006); and although children encounter many instances of dissolving in everyday life (cocoa, sweets, soap), these may be less frequent than instances of sinking/absorption. Associative learning accounts of causal induction from everyday experience would also predict lower levels of learning for less temporally contingent events even if these were encountered with equal frequency, since the contingency would be less evident.

Future research could directly test whether the speed of a process affects children's ability to track it during observation – consistent with Maurice-Naville and Montangero's (1992) reports of elementary school children's difficulties grasping the effects of slowly progressing forest disease – and whether longer timeframes demand more concentration – consistent with Rieber's (1991) finding that fourth graders only benefited from an animated demonstration of Newton's second law when this was presented in short sequences.

Differences between components of causal task responses

Results demonstrated a clear progression with age on all causal components. The different components differed widely in difficulty, however, with description ahead of prior knowledge, and explanation trailing far behind. Just under 80% of children across all ages and tasks made and reported completely accurate observations; even for the most difficult dissolving process, 31% of the youngest children achieved the top score for description. For prior knowledge, nearly 58% of children got the highest scores overall. The lack of progress between Y3 and Y5 on these two components therefore seems largely attributable to good levels of performance having already been achieved by 8-year-olds. Since children did well on description, they plainly have the language for observation, and if prediction lags behind, it seems more likely to be due to lack of knowledge than lack of words.

Causal explanation scores trailed by long way, with only about 9% reaching top scores across the three tasks and all age groups, and mechanism responses were comparatively unusual – only a fifth of children, the majority in the oldest age group, identified any underlying causal mechanism – although once that insight was achieved, it appeared to be quickly extended to multiple processes. Children's explanations were typically limited to abstracting causal variables from the observed objects, and the youngest age group was commonly unable even to explicitly relate causal factors to the observed speed differences, that is to actually treat these as variables. Variable abstraction implies increased selective attention to the perceptual input across observed instances of processes, but identification of mechanisms to account for these variables imposes substantially greater inferential demands. For causal explanation, therefore, the lack of progress between Y3 and Y5 is attributable to children finding it difficult to move on from abstraction of variables to identification of mechanisms.

One can suggest that children's mechanism responses might have been the product of instruction rather inference. However, note that four points indicate this is on the whole unlikely: (1) the relatively weak relationship of prior knowledge to explanation, compared to that between description and explanation, suggests responses were constructed largely from observation; (2) although they were infrequent, mechanism responses were given by some younger participants, who would not yet have had more technical instruction in concepts like density, porosity, and solubility; (3) conversely, despite the fact that the topics were within the curriculum content, mechanism responses were only given by a minority of Y5 children, and there were clear differences between sinking, absorption, and solution, again suggesting that observing the contrast phenomena was more important – the difficulties in perceiving the solution events are a more plausible source of these variations; and (4) provision of mechanism responses was influenced by non-verbal ability, while effects of instruction seem more likely to be mediated by verbal ability.

However, if observation was central to children's performance, there is plainly a long way from good descriptive ability to higher-level inferential responses. The striking lag in the explanation component may be because explanation of mechanisms places higher demands on children's language. We know from the preschool studies involving distinct causal systems that in principle children of this age have no limitation in thinking about mechanism (Buchanan & Sobel, 2011; Bullock *et al.*, 1982; Schlottmann, 1999; Shultz, 1982). Further investigation is therefore needed to establish how far prompting children to give thorough explanations mitigated the verbal demands of our tasks. However, as we discuss below, there is good reason to think that language skills *per se* were not the issue. It seems more likely that children find continuous processes hard to analyse at this level.

Comparisons with adults using our paradigm are needed, since mechanism reports may be sparse even among them; no research has examined this to date.

Do verbal and non-verbal ability predict children's causal reasoning about continuous processes?

Verbal ability, as indexed by vocabulary, did not predict performance on any of the three causal components, or mechanism responses, arguably the most verbally demanding level of inference. This might potentially reflect the broadly homogenous nature of the sample, with relatively high socioeconomic background – except that vocabulary scores were normally distributed, and there was no indication that variance in verbal ability was limited enough to prevent it having predictive power. We note also that the verbal task we used was generic, though, with a focus on everyday language. During testing, we observed that children who reported causal mechanisms tended to use more abstract/scientific vocabulary. Such language may be more specifically related to causal reasoning and could therefore be more discriminating. Dündar-Coecke & Tolmie, (2019b) present preliminary data suggesting that the relationship is real, though no investigation has focused on this previously, again, suggesting a gap in the literature.

Although it is not entirely clear whether grasp of causal processes is influenced by verbal competence, on these data its impact appears to be limited. One possibility is that non-verbal impressions from previously and currently observed phenomena are more critical than explicit concepts, and that these impressions are combined and then *translated* as far as possible into the verbal domain (see Dündar-Coecke & Tolmie, in press, for a discussion) – meaning that a certain level of verbal ability is required to get this off the ground, in line with the mediation effect in the path model for mechanism. Scientific vocabulary may assist further with this process of translation.

In contrast to verbal ability, non-verbal ability as indexed by block design uniquely predicted all components of the causal tasks, including inference of mechanism. However, the logarithmic relationship of block scores to causal scores reflects a steeper gradient of relationship at lower levels of performance on the latter, indicating that this is where generic non-verbal ability made most difference. This is consistent with the lower explained variance in the regression model for the highest level of explanation response, mechanism, suggesting that at this level in particular, some additional non-verbal factor might be a key element in children's cognition of continuous causal processes.

The strong influence of non-verbal ability on causal reasoning about continuous processes in words may seem surprising at first. However, to go beyond the observation of a continuous process, as is necessary for thinking about causal variables and connecting mechanisms, requires mental imagery, combining both visible (objects) and intervening invisible (e.g., density, buoyancy) elements. The role of non-verbal ability in this kind of thinking has been explored in another study (Dündar-Coecke & Tolmie, 2019b), which also found that higher levels of reasoning about causal effects do not rely primarily on verbal but on non-verbal competences, suggesting that the finding here is robust. These competences may further include detailed analysis of the spatial–temporal characteristics of processes: the ability to extract key dimensions of information from object states that change over time, to conceive of the sequence of dynamic transformations that underlie such observed change, and to project these transformations onto past, present, and future experiences. This segmentation of a continuous observation into meaningful steps is potentially the additional element suggested above. A study in progress evaluates this view further.

Conclusion

This is the first research on children's causal reasoning in words about continuous physical processes, including their inferences of the invisible mechanisms that underlie these. Even the youngest participants were able to give sensible descriptions of what they could observe, but children's causal explanations, both their abstraction of causal variables and inferences of mechanism, developed more slowly than their observation skills and application of prior knowledge, suggesting that thinking about continuous natural processes may be hard for elementary school children. However, although this ability lags behind reasoning about mechanism in more transparent contexts, and is affected by the nature of the processes involved, it is nevertheless clearly within their competence. We have also illustrated that – in this sample at least – non-verbal ability is a clear predictor of children's reasoning.

Research on causal cognition should encompass not just analysis of events and machines with distinct cause–effect segments, as studied throughout psychology, but also analysis of continuous processes. Such analysis is within the reach of elementary school children, and we need to study it more if we wish to understand how they make the transition towards adult – and genuinely scientific – causal reasoning.

Acknowledgements

This research was funded by an internal grant provided by the Institute of Education, University College London, and by an award from UK Economic and Social Research Council to the first author.

References

Buchanan, D. W., & Sobel, D. M. (2011). Mechanism-based causal reasoning in young children. *Child Development*, *82*, 2053–2066. https://doi.org/10.1111/j.1467-8624.2011.01646.x

Buchsbaum, D., Griffiths, T. L., Plunkett, D., Gopnik, A., & Baldwin, D. (2015). Inferring action structure and causal relationships in continuous sequences of human action. *Cognitive Psychology*, *76*, 30–77. https://doi.org/10.1016/j.cogpsych.2014.10.001

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 209–254). New York, NY: Academic Press.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Dündar-Coecke, S., & Tolmie, A. (in press). A short term intervention improved children's insights into causal processes. *The Psychology of Education Review*.

Dündar-Coecke, S., & Tolmie, A. (2019a). What makes causal mechanism more accessible across development?

Dündar-Coecke, S., & Tolmie, A. (2019b). Non-verbal ability and scientific vocabulary are more robust predictors of causal reasoning in science than generic language. *Mind Brain and Education*.

Danks, D. (2005). Causal learning from observations and manipulations. In M. Lovett & P. Shah (Eds.), *Thinking with data*. Hillsdale, NJ: Erlbaum.

Gelman, S. A., & Kremer, K. E. (1991). Understanding natural cause: Children's explanations of how objects and their properties originate. *Child Development*, *62*, 396–414. https://doi.org/10.2307/1131012

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment? *Psychological Review*, *62*, 32–41. https://doi.org/10.1037/h0048826

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, *111*(1), 3–32. https://doi.org/10.1037/0033-295X.111.1.3

Gopnik, A., & Meltzoff, A. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, *111*, 3–32. https://doi.org/10.1037/0033-295x.111.1.3

Kitcher, P. (1989). Explanatory unification and the causal structure of the world. In P. Kitcher & W. Salmon (Eds.), *Minnesota studies in the philosophy of science. Vol. 13, Scientific explanation* (pp. 410–505). Minneapolis, MN: University of Minnesota Press.

Lagnado, D., & Sloman, S. A. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 856–876. https://dx.doi.org/10.1037/0278-7393.30.4.856

Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S. A. (2007). Beyond covariation: Cues to causal structure. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 154–172). Oxford, UK: Oxford University Press.

Legare, C. H. (2014). The contributions of explanation and exploration to children's scientific reasoning. *Child Development Perspectives*, *8*, 101–106. https://doi.org/10.1111/cdep.12070

Legare, C. H., & Clegg, J. M. (2014). *The development of children's causal explanations*. Retrieved from https://static1.squarespace.com/static/53485734e4b0fffc0dcc64c2/t/53ac13e4e4b0e1ede09d0031/1403786212544/the-development-of-childrens-causal-explanations++.pdf

Legare, C. H., & Lombrozo, T. (2014). Selective effects of explanation on learning during early childhood. *Journal of Experimental Child Psychology*, *126*, 198–212. https://doi.org/10.1016/j.jecp.2014.03.001

Liu, X., & Lesniak, K. (2006). Progression in children's understanding of the matter concept from elementary to high school. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *43*, 320–347. https://doi.org/10.1002/(ISSN)1098-2736

Maurice-Naville, D., & Montangero, J. (1992). The development of diachronic thinking: 8–12-year-old children's understanding of the evolution of forest disease. *British Journal of Developmental Psychology*, *10*, 365–383. https://doi.org/10.1111/j.2044-835X.1992.tb00583.x

Michotte, A. E. (1946/1963). *The perception of causality*. New York, NY: Basic Books.

Piaget, J. (1929). *The child's conception of the world*. New York, NY: Harcourt, Brace Jovanovich.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64–99). New York, NY: Appleton-Century-Crofts.

Rieber, L. P. (1991). Effects of visual grouping strategies of computer-animated presentations on selective attention in science. *Educational Technology Research and Development*, *39*, 5–15. https://doi.org/10.1007/BF02296567

Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton, NJ: Princeton University Press.

Schlottmann, A. (1999). Seeing it happen and knowing how it works: How children understand the relation between perceptual causality and underlying mechanism. *Developmental Psychology*, *35*, 303–317. https://doi.org/10.1037/0012-1649.35.1.303

Schlottmann, A., Ray, E. D., Mitchell, A., & Demetriou, N. (2006). Perceived physical and social causality in animated motions: Spontaneous reports and ratings. *Acta Psychologica*, *123*, 112–143. https://doi.org/10.1016/j.actpsy.2006.05.006

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology*, *103*, 759–775. https://doi.org/10.1037/a0025140

Shanks, D. R., & Dickinson, A. (1987). Associative accounts of causality judgment. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 21, pp. 229–261). New York, NY: Academic Press.

Shultz, T. R. (1982). Rules of causal attribution. *Monographs of the Society for Research in Child Development*, *47*, 1–51. (Serial No. 194).

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22. https://doi.org/10.1037/0033-2909.119.1.3

Sloman, S. A., & Lagnado, D. (2005). Do we "do"? *Cognitive Science*, *29*, 5–39. https://doi.org/10.1207/s15516709cog2901_2

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E.-J., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, *27*, 453–489. https://doi.org/10.1207/s15516709cog2703_6

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science*, *12*, 49–100. https://doi.org/10.1207/s15516709cog1201_2

Tolmie, A., & Dündar-Coecke, S. (in press). Lifespan conceptual development in science: Brain and behaviour. In M. S. C. Thomas, D. Mareschal & I. Dumontheil (Eds.), *Educational neuroscience: Development across the life span*.

Vosniadou, S. (2014). Examining cognitive development from a conceptual change point of view: The framework theory approach. *European Journal of Developmental Psychology*, *11*, 1–17. https://doi.org/10.1080/17405629.2014.921153

Waldmann, M. R., & Hagmayer, Y. (2005). Seeing versus doing: Two modes of accessing causal knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 216–227. https://doi.org/10.1037/0278-7393.31.2.216

Walker, C. M., Lombrozo, T., Williams, J. J., Rafferty, A. N., & Gopnik, A. (2017). Explaining constrains causal learning in childhood. *Child Development*, *88*, 229–246. https://doi.org/10.1111/cdev.12590

Wechsler, D. (2011). *Wechsler abbreviated scale of intelligence, second edition (WASI-II)*. San Antonio, TX: NCS Pearson Education.

Wellman, H. M., & Gelman, S. A. (1998). Knowledge acquisition in foundational domains. In D. Kuhn & R. Siegler (Eds.), *Cognition, perception, and language. Vol. 2 of the handbook of child psychology* (5th ed., pp. 523–573). Editor-in- Chief: William Damon. New York, NY: Wiley.

White, P. (2014). Singular clues to causality and their use in human causal judgment. *Cognitive Science*, *38*, 38–75. https://doi.org/10.1111/cogs.12075

Wilkening, F., & Cacchione, T. (2011). Children's intuitive physics. In U. Goswami (Ed.), *The Wiley–Blackwell handbook of childhood cognitive development* (2nd ed., pp. 473–496). Oxford, UK: Wiley-Blackwell.

Zaitchik, D., Solomon, G. E. A., Tardiff, N., & Bascandziev, I. (2016). Conceptual change: Where domain-specific learning mechanisms meet domain-general cognitive resources. In D. Barner & A. S. Baron (Eds.), *Core knowledge and conceptual change* (pp. 73–88). Oxford, UK: Oxford University Press.

Received 30 December 2018; revised version received 7 March 2019

Appendix: Examples of explanation responses

Phenomena	Level 1	Level 2	Level 3
Sinking	'They are heavy and they sank to the bottom'	'The stone is heavier than the berry so they sank to the bottom differently'	'They are both heavier than the water and cannot hold air in it so they sank to the bottom. But the stone sank quicker than the berry because it's got more stuff in it so the water can't hold it up as it did to berry'
Absorption	'If you dip the paper in the water they get wet because they're soft'	'The tissue paper is thinner than the other paper so water rises faster in it'	'The tissue paper has holes in it that help water to rise up. Water holds on the walls of the holes and layers and that helps it to climb up. Other paper has some space in it, but not as much as the tissue paper'
Solution	'They go into water because they're small and spread out'	'The table salt is smaller than the rock salt so it disappears quicker'	'The size of the two types of salt is different. And this is more rocky so water cannot go into it easily. They both dissolve in the water, but rocky one takes more time than the table salt'