Using dialogic interventions to decrease children's use of inappropriate teleological explanations

Jonathan Grant Halls, Shaaron Elizabeth Ainsworth and Mary Colette Oliver

Learning Science Research Institute, School of Education, University of Nottingham, Nottingham, United Kingdom

A belief in teleology is often suggested to be a barrier in children's science education. Many studies have catalogued children's use of teleological explanations, but fewer have developed approaches to tackle children's use in scientific contexts. This paper reports two studies that utilised dialogic interventions alongside Concept Cartoons to do just that. Study 1 (5- to 7-year-olds, n = 54) addressed teleological explanations for natural phenomena (e.g., snow or rainbows) and Study 2, (9- to 10-year-olds, n = 24) explored organisms' traits (e.g., giraffes' necks or zebras' stripes). Both studies found that after only short discussions about styles of explanation in science, children's acceptance of teleological explanations was significantly reduced and they were more likely to endorse appropriate scientific explanations. These results suggest that teleology need not be a major barrier to teaching and learning about causality as it can quickly and effectively be addressed.

Introduction

A considerable body of work suggests that children, aged 3- to 10-years-old, primarily use teleological explanations to explain the existence of natural kinds (e.g., Kampourakis, Pavlidi, Papadopoulou & Palaiokrassa, 2012; Kelemen, 1999). Such explanations usually propose a purpose (e.g., muddy puddles exist *for* children to jump in) or suggest goal-driven behaviour (e.g., tigers *decided* to grow stripes for better camouflage) as a rationale for why something exists. Existing literature suggests children's use of teleology is prevalent. For example, Kelemen (1999) found that children endorsed teleological explanation ~70% of the time for the existence of natural organisms and ~85% of the time for natural objects. In science classrooms, children's preconceptions about different scientific theories can be problematic, as they are argued to be difficult to change (Holman & Yeomans, 2018). Teleological preconceptions, specifically, are suggested to be a significant barrier when teaching evolutionary concepts (Kampourakis, 2014; Trommler & Hammann, 2020). To date, most research focuses on cataloguing preconceptions, with less having explored how to limit children's use of or develop their understanding of the limitations of this explanation. This paper directly addresses this by reporting two studies designed to limit children's endorsement of teleology by developing their understanding of appropriate scientific explanations.

These studies are underpinned by the notion that teleology is a collection of different explanatory styles rather than a method of reasoning. While teleological explanations will be based on some form of reasoning process, they may not always be the product of explicit teleological reasoning, i.e., a design stance, whereby everything exists due to some form of design to meet a goal or desire. This is illustrated by considering different forms of teleology. Types of teleology include, but are not limited to: design-teleology, something is explained as having been created explicitly by an external or internal agent for a specific purpose (e.g., rain is for watering plants) (Kampourakis, 2020); functional-teleology, something is explained with reference to the function it provides for an organism (e.g., anteaters' claws are for digging) (Zohar and Ginossar, 1998); or relational-teleology, something is explained in relation to a purpose for which it appears to be used or an effect that it causes, typically in connection to humankind (e.g., night is a time that people use for sleeping or day is for going to school) (Halls, Ainsworth and Oliver, 2018). Such relational-teleological explanations reflect individuals' experiences of the world. This is what ojalehto, Waxman and Medin (2013) term a perspectival relationship; how living things are connected to their environment. These explanations result from reasoning about relationships in the natural world rather than reasoning about design or purpose in the world. Viewing teleology as a style of explanation that may or may not be due to an explicit teleological reasoning process has several implications, including the interplay between cultural patterns of language and children's use of teleological explanations, and the impact that this has on designing methods of instruction. Drawing on suggestions by Trommler and Hammann (2020) and Halls, Ainsworth and Oliver (2018), the work outlined here designs interventions that support children to recognise incorrect teleological explanations in order to 'chip away' at fragments of knowledge that may cause children to endorse such explanations in scientific contexts.

Understanding teleology

Learning about scientific concepts is complicated once children form preconceived ideas, acting as barriers to conceptual change (diSessa, 2014; Holman & Yeomans, 2018). While some studies have taught the correct scientific concept to counter children's use of teleological explanations (e.g., Kampourakis & Zogza, 2009; Kelemen, Emmons, Schillaci & Ganea, 2014), we have found no research on specifically fostering understanding of the appropriateness of teleological explanations. Instead, most studies have taken snapshots of when children use teleological explanations.

Barnes, Evans, Hazel, Brownell and Nesse (2017) found that students may still endorse incorrect teleological explanations after explicit instruction on evolution. They propose tacking the use of teleology alongside explicit instruction on scientific concepts. We suggest this could be achieved through critical discussion about teleological styles of explanation. Rigorous appraisal of ideas enables children to develop their scientific literacy (Henderson, MacPherson, Osborne & Wild, 2015), which includes understanding appropriate explanations. Consequently, interventions highlighting the inappropriateness of teleological explanations may be effective at limiting children's use (Trommler & Hammann, 2020).

Halls et al. (2018) argued that children's use of teleological explanations may result from fragmented knowledge about the natural world. Using structured interviews, children (5- to 8-years-old) were asked to explain the existence of different natural phenomena (time-constrained events or processes, such as weather) to explore the effect of different question wordings and the question focus (e.g., rain or snow). Results suggested that children's teleological explanations varied depending upon the natural phenomena they were asked to account for. For example, discussing the reasons for waterfalls led children to generate significantly more teleological explanations than rainbows. Furthermore, there was variation in the type of teleological explanation used, with some natural phenomena leading to more design-teleological explanations but the majority associated with more relational-teleological explanations. diSessa (2014) argued that children have a fragmented knowledge structure, which leads to the hope that children's preconceptions could be altered by educators reframing children's understanding of where and when teleological explanations are appropriate.

While prior work establishes that children use high levels of teleological explanation (e.g., Kelemen, 1999), there is little research on explicitly countering this. In one example (Bartov, 1978), 10th-grade students took part in a series of biology lessons that explored the causes of organism action. Learning activities included hypothesising about both teleological causes (such activities happen because they are useful for the organism) and scientific accounts and then experimentally evaluating these hypotheses. Following instruction, students were better able to distinguish between causal and teleological explanations. A similar study with high school and undergraduate students (Richardson, 1990) focused on teleological and mechanistic explanations for bodily functions. Following a lecture on the difference between such explanations, on average, students selected teleological explanations on 12% of the time. Without instruction, students chose teleological explanations for 61% of the test items, suggesting that such short instructive activities are effective. However, the study's teleological explanations are arguably a form of acceptable functional-teleology (Trommler, Gresch & Hammann, 2017). Both studies show that students can be taught to recognise and reject teleological explanations; however, they were conducted with older students, long after a preference for teleological explanations has been established.

Kelemen and colleagues (Kelemen, Emmons, Seston Schillaci & Ganea, 2014; Emmons, Smith & Kelemen, 2016) conducted studies with younger children (5- to 8-year-olds) using a storybook approach. The book detailed the gradual adaptation of a species with trait variation over time, focusing on the diminishing of the disadvantageous trait form and the advantageous trait form's proliferation. After reading the book, children showed enhanced understanding of natural selection. These studies show that young children are capable of understanding evolutionary mechanisms. However, the focus was on teaching correct scientific accounts rather than supporting children in recognising incorrect teleological explanations. Both may be required to tackle children’s use of teleology.

Purpose of studies

While there has been some work on recognising and rejecting teleological explanations, there is still a considerable imbalance between work on documenting teleology and work on countering teleology. This imbalance led to the two studies reported here. Both explore if discursive activities about the nature of explanation in science support children in recognising and understanding why teleological explanations can be inappropriate. The aim is to limit children's endorsement of teleological explanations, likely a crucial first step in reducing children's generation of teleological explanations. The studies focused on two different age ranges (5- to 7-year-olds and 8- to 9-year-olds) and two different areas where teleological explanations are often used (natural phenomena and organisms' traits).

Study 1

Study 1 explored the impact of dialogic activities on children's understanding of why causal explanations are appropriate and teleological explanations are inappropriate to account for natural phenomena.

Method

Intervention design

Intervention underpinning

Discourse is integral to developing conceptual understanding, as language is a tool to explain, share and analyse ideas (Mercer & Howe, 2012). Discourse-based science activities are beneficial when word meaning can cause confusion (Loxley, Dawes, Nicholls & Dore, 2010). These forms of discussions, where meanings are clarified and individuals develop a shared understanding of expressions, are pertinent to the design of the intervention. For example, in Study 1, there was a need for children to differentiate between design-teleology and relational-teleology for the existence of natural phenomena. While explanations about natural phenomena could seem to suggest *X is for Y*, what is meant is that *X can be used to do Y*.

Dialogic teaching is a pedagogic approach that supports discussion-based activities in science (Grugeon, Hubbard, Smith & Dawes, 2012). It involves collaborative, critical exploration of ideas, in a teacher-guided group setting where all contributions are allowed and related to personal experience, ideas are questioned, statements are justified, and reflection is encouraged (Alexander, 2006). The underpinning idea is that dialogue is a driving force behind individuals' critical consideration of conceptions, and ultimately, conceptual change. Research on dialogic activities shows improvements in children's performance on science assessments (Mercer, Dawes, Wegrif & Sams, 2004) and benefits in numerous other areas (Howe & Abedin, 2012).

Dialogic teaching explores an issue in-depth, enabling different views to be considered and shared understanding to be established (Grugeon et al., 2012). This is valuable when social or cultural factors could influence children's preconceptions and thoughts about ideas under discussion. The capacity of dialogic teaching to foster conceptual change made it a beneficial theoretical underpinning for the designed interventions. Concept Cartoons were used as practical resources to enact the principles of dialogic teaching in the designed intervention.

Concept Cartoons are effective at generating discussion about complicated scientific concepts (Chin & Teou, 2009; Naylor & Keogh, 1999). They typically consist of a cartoon depicting a certain scientific concept (e.g., causality or material properties) with several characters offering explanations about the science. For example, one shows a snowman with children offering statements such as '*Don't put the coat on the snowman. It will melt him*' or '*It will keep him cold and stop him melting'*. These contradictory statements form the starting point for discussion.

Intervention structure

Table 1 details the overall structure of the intervention used for both studies' treatment conditions; study-specific alternations are detailed where appropriate.

Table 1

Intervention structure for the treatment conditions

Stage	Description
Introduction (1-2 minutes):	Children told discussion would focus on a certain topic (e.g., clouds or rabbits' ears). Specifically focusing on why it exists, and what constitutes a good or silly scientific explanation.
Discussion activity (10-12 minutes):	Children were shown a Concept Cartoon for the discussion topic, with four statements about its existence. These were a mix of teleological and scientific accounts. Firstly, children discussed if individual statements did (good) or did not (silly) account for the existence of the discussion topic. Secondly, children reflected on the commonalities between good or silly explanations.
Plenary (4-5 minutes):	Additional teleological and scientific explanations were presented and children were asked why it was a good or silly explanation.

Assessment

Children were assessed using Concept Cartoon style prompts, which consisted of a single teleological or scientific statement to evaluate. There were two stages:

1. Children judge the explanation as good (appropriate) or silly (inappropriate), allowing for assessment of children's initial judgements of explanations;

2. Children provide a rationale for their judgement, allowing for assessment of children's understanding of the elements of appropriate/ inappropriate explanations.

Explanation judgement tasks (stage one) are typically used in studies on children's use of teleology (Trommler & Hammann, 2020) but struggle to capture the nuance of children's reasoning (Trommler et al., 2017). Therefore, asking for rationales of judgements (stage two) aimed to measure children's reasoning behind their judgements.

Design

A 2x2 mixed experimental design was used, the within-groups variable was time (pre-test and post-test), the between-groups variable was condition (intervention and control). A no-treatment control was used, as there was no comparable treatment to the intervention.

Participants

Fifty-three children, aged 5- to 7-year-old, took part from two schools. Both schools were single-form entry primary schools in a large city in the East Midlands, UK. Due to school policy, children's date of births were not provided by one of the schools; therefore, analysis by children's age was not conducted. Twenty-five children (5-to 6-years-old) took part from school one (Female = 9) and 28 children (6- to 7-years old) took part from school two (mean age = 84.5 months old, *SD* = 9.44, Female = 11). In each school, children were randomly assigned to the treatment (27 children, School One = 13, School Two = 14, Female =14) or no-treatment control condition (26 children, School One = 12, School Two = 14, Female = 6).

Treatment

Treatment condition participants were randomised into different groups (three to four members) every session. This was designed to reduce the effect of having consistent group members, thereby creating a wider selection of viewpoints within groups, across the sessions. The treatment sessions were based on the generic intervention structure outlined in Table 1. Four sessions covered the topics of clouds, rain, rainbows and waves. Each included an introduction where children were presented with a Concept Cartoon (Concept Cartoon statements are outlined in Appendix 1), followed by a discussion concerning whether the statements were good scientific explanations for why X (e.g., rain) exists. In the plenary, children were asked to quickly judge additional statements as good or bad and provide a rationale for their judgement to the group. During the discussion and plenary, children were prompted by the researcher with talk moves. These were based on those suggested by Michaels and O'Conner (2012) and were used to enact the principles of dialogic teaching set out by Alexander (2006).

Talk moves fell into four areas. Firstly, supporting children to share or expand upon ideas, which included reminding them to reflect on Concept Cartoon statements, asking children to elaborate on their points (e.g., "can you tell us more about that?") or rephrasing children's comments for clarity (e.g., "so you are saying …"). Secondly, encouraging children to add to or challenge others' comments. These included asking children if they agreed with a comment by another child and why (e.g., "does everyone think the same as Child A"), encouraging deeper reflection on a comment (e.g., "can anyone add anything else to what Child A said?"), highlighting differences and similarities in opinions (e.g., "Child B, you said something similar to Child A earlier, do you think you meant the same thing? What do we think is different about what Child A and Child B said?"). Thirdly, prompting children to support their ideas with evidence, through highlighting discrepancies (e.g., "has everyone had the same experience as Child A?", "Child B, is what you said always true?", "can we think of times when this is not true?") or requests about reasoning (e.g., "so what makes you think that?", "is there anything that's happened to you to make you think that way?"). Finally, talk moves were used to focus on the main content (the appropriateness of statements) and were used when the conversation became unrelated to session content (such as digressions about previous snowball fights) or when children were not evaluating the statement in terms of how it accounts for the existence of X. Regarding the latter, at times children would focus on how pleasurable or 'good' the effect of the statement was. For example, in 'rain is for watering the plants', children may discuss how it is good that plants get water, therefore, prompts were used to focus the discussion (e.g., "yes that's true, people often say that rain helps to water the plants, but does it help us explain why rain happens?").

Assessment

Participants were informed that 'I have some cartoons about different topics and all I want you to do is to tell me if they are a good or a silly explanation for why something exists'. Importantly, the natural phenomena assessed were different to those in the intervention. Therefore, the assessment was not evaluating if children could remember a correct answer, rather investigating if children could apply what they had learnt to new examples of natural phenomena. All participants were assessed using twelve statements across four topics (day, snow, lightning and wind). For example, for snow, children were presented with statements suggesting there is snow for telling some animals to hibernate over winter (design-teleology), so children can make snowmen (relational-teleology) and because it is really cold and the raindrops freeze and turn into snowflakes (causal). These topics were based on Halls et al. (2018) and were those with the highest proportion of teleological explanations. The full set is detailed in Appendix 1.

Measures

Judgements were coded as correct (e.g., teleological statement = silly) or incorrect (e.g., scientific statement = silly). Rationales were coded using the rubric in Table 2. Second blind coding of 10% of the data (126 responses) indicated a good level of agreement ($K = .787$, $p < .001$).

Table 2

Study 1 rubric for rationales

Score	Explanation type	Example
1	The rationale discusses the style of explanation used in the statement. That is, the statement is claimed to be good because it offers a causal explanation of why/how X happens. The statement is silly because it does not offer a causal explanation to why/how X happens. Instead, it states what X is used for Y or states X is for Y when it is not for anything	[There is snow because of rain freezing] is good cos it tells you how the snow happens. [Wind is for flying kites is silly] because it is just telling you what children can do, not why there is wind
0	Any other form of rationale, including incorrect responses, non-response or axiomatic justification.	[Snow is for animals to hibernate is good] because there is snow to tell animals when they need to go to sleep. [There is snow because of rain freezing] is good cos then people can have snowball fights

Procedure

Data collection in both schools took place across seven weeks (School, One March to April; School Two, June to July). The timetable for the treatment group was the same in both schools:

- Week one, pre-test – individual assessment;
- Week two, treatment session 1 – clouds;
- Week three, treatment session 2 – rain;
- Week four, break – school half-term;
- Week five, treatment session 3 – rainbows;
- Week six, treatment session 4 – waves;

- Week seven, post-test – individual assessment.

The no-treatment control group only took part in the week one and seven tests. For assessment and treatment sessions, children were removed from their classroom while the rest of the class continued with lessons unrelated to the intervention. Pre- and post-test individual assessments lasted approximately seven minutes; group treatment sessions lasted approximately 15 minutes. All school activities were conducted by one of the researchers, who has a background as an early years teacher.

Results

For the following analyses, two variables were created: judgement score and rationale score. These were the participants' mean score for their twelve judgements or twelve rationales (maximum = 12). Before testing the hypotheses, two one-way ANOVAs were conducted to evaluate pre-existing differences between schools or age groups. The independent variable was group, with four levels (School one control, School one treatment, School two control, School two treatment), the dependent variables were judgement or rationale score at pre-test. The four groups' judgement scores were not significantly different, F (3, 52) = 1.83, $p.$ = .154, η^2 = .101, nor were there rationale scores, F (3, 52) = 2.67, p = .057, η^2 = .141. Consequently, the following analyses combined the treatment (School one and two treatment groups) and control (School one and two control groups) groups.

Analysing children's judgements

A 2x2 mixed ANOVA was conducted to investigate the intervention's influence on children's judgement of statements. The within-groups independent variable was time (pre-test and post-test), the between-groups independent variable was condition (control or treatment) and the dependent variable was judgement score.

The analysis showed a main effect of time (pre-test vs. post-test), *F* (1, 51) = 43.92, *p*. < .001, η_p^2 = .463, of condition (control vs. treatment), *F* (1, 51) = 12.32, *p*. < .001, η_p^2 = .195, and a significant interaction between time and condition, *F* (1, 51) = 103.45, *p*. < .001, η_p^2 = .670. Consequently, multiple pairwise comparisons were conducted, with Bonferroni corrections. These showed no difference between the pre-test scores, *t* (51) = 1.20, *p*. = .932, *d* = .33, but a significant difference between post-test scores, *t* (51) = 7.35, *p*. < .001, *d* = 2.02. Paired *t*-tests showed a significant decrease in children's correct endorsement of explanations from pre- to post-test for the control condition, *t* (25) = -3.09, *p*. = .019, d_{rm} = 0.38, and a significant increase for the treatment condition, *t* (26) = 10.35, *p*. < .001, d_{rm} = 2.08. Table 2 shows the mean judgment scores by condition.

Table 3

Study 1 mean judgement and rationale scores, by time and condition (Max. score = 12)

	Mean judgment score	
	Control	Treatment
Pre-test score (*SD*)	6.04 (2.34)	5.26 (2.36)
Post-test score (*SD*)	5.04 (2.71)	10.00 (2.16)
	Mean rationale score	
	Control	Treatment
Pre-test score (*SD*)	2.38 (3.44)	2.07 (3.30)
Post-test score (*SD*)	2.03 (3.54)	7.70 (4.05)

Analysing children's rationales

A 2x2 mixed ANOVA was conducted to investigate the influence of the intervention on children's understanding of the elements of appropriate, or inappropriate, explanations about the existence of natural phenomena. Independent variables remained the same, the dependent variable was participants' rationale score.

The analysis showed a significant main effect of time, F (1, 51) = 41.70, $p. < .001$, η_p^2 = .450, condition, F (1, 51) = 8.83, $p. = .004$, η_p^2 = .148, and a significant interaction between time and condition, F (1, 51) = 53.34, $p. < .001$, η_p^2 = .511. Bonferroni corrected independent *t*-tests showed no difference between conditions at pre-test, t (51) = 0.33, $p. > .999$, d = .09, but a significant difference at post-test, t (51) = 5.41, $p. < .001$, d = 1.48. Paired *t*-tests showed no change in the control condition t (25) = -1.08, $p. > .999$, d_{rm} = 0.09, but a improvement in the treatment condition, t (26) = 7.58, $p. < .001$, d_{rm} = 1.50, from pre- to post-test with children providing more rationales that demonstrated awareness of appropriate or inappropriate explanations. Table 2 shows the mean rationale scores by condition.

These rationales suggest a developing understanding that appropriate scientific explanations are causal (e.g., the explanation must suggest how something happens) and that teleological explanations are inappropriate (e.g., natural phenomena do not exist for a specific purpose, even though people may refer to natural phenomena appearing to be used for something). Correct rationales about scientific explanations at post-test from children in the treatment condition included: *There is day because the Earth is spinning around and when the bit we are on faces the sun, then we have daytime*[1] is good "because it tells you why we have daytime, it's when the sun is shining on us", and *There is lightning because parts of clouds bump together and make electrical charge and when the cloud gets full of electric charge a lightning bolt comes out* is good "cos it is telling us how the lightning is happening, cos all the clouds bump together". Examples of rationales that disregarded teleological statements included: *There is snow*

[1] Italics in examples indicate Concept Cartoon statements

for telling some animals to hibernate over winter is silly because "hibernate is like a science word but it doesn't tell us how [snow] is generated"; *There is wind for blowing the leaves of trees* is silly because "leafs [sic] get blown off when the wind has to come, but it is not telling us about the wind"; *There is day so children can go to school* is silly because "it is not telling us how it is made, it is just telling up about children going to school"; *There is snow so children can make snowmen* is silly "because you might want to make snowmen but it is not for that and it's not telling you how it is made".

For comparison, the following are examples of incorrect pre-test rationales about lightning. *There is lightning for lighting up a dark, stormy sky* is good ", so when you need a toilet you can see where it is." *There is lightning so people can use it to power electric appliances, like the TV or lights* is good "so people can watch TV." *There is lightning because parts of clouds bump together and make electrical charge, and when the cloud gets full of electric charge, a lightning bolt comes out* is silly "cos people can get hit by it".

Study 1 summary

The results show the intervention was successful. Young children's awareness of the appropriateness of different explanatory styles was improved through dialogic activities. For participants in the treatment condition, their average score for correct judgements about the existence of natural phenomena increased substantially to over 80%, whereas those control condition showed little change. Thus, after the intervention, there was a significant improvement in the intervention group's judgements of explanations, even though the topics had not been covered during the four teaching sessions. Regarding children's rationales for their judgements, again the control group's average score showed little change. However, the participants in the treatment condition improved considerably.

The results indicate that children's teleological preconceptions were not inherently difficult to change. Their competence in evaluating explanations dramatically improved after short dialogic activities that lasted approximately 40 to 50 minutes. From this study, it appears that children are much more competent than usually assumed as, after support, they can recognise and disregard teleological explanations from a young age. As the study did not use delayed post-test assessment, further work is required to explore how children retain this ability over time.

Study 2

In Study 1, it can easily be argued that teleological explanations about natural phenomena are objectively scientifically inappropriate. However, when considering the debate about whether explanations concerning the development of organisms' traits can make use of teleology is more complex. It is a complicated debate, but opposite and distinctive perspectives can be identified concerning the appropriateness of teleological explanations for organisms' traits (e.g., tigers' stripes are for camouflage). One perspective is that functional-teleological explanations are always inappropriate (see Galli & Meinardi, 2011; Hanke, 2004), as their use could result in: 1) organisms' desires being used as the cause of evolutionary change; 2) evolution being thought of has having a predetermined path or goal; 3) accounting for the existence of organisms' traits solely in terms of their function.

The other position is that teleological exposition is a convenient and natural style of explanation (Ruse, 1989). Explanations for organisms' traits, which focus on the function and purpose of traits are understandable, as biological traits often cannot be explained without reference to the function they provide for an organism (Trommler &

Hammann, 2020). Furthermore, studies show that children can use teleological explanations as a learning heuristic without sacrificing scientific rigour (Zohar & Ginossar, 1998); essentially using teleology as an additional method of description rather than a stand-alone explanation.

It would appear that these two arguments have different origins, and that confusion may result from the overarching term *teleological explanation*. Therefore, one solution, also proposed by Trommler and Hammann (2020), is to take a more nuanced perspective on teleology and considers sub-types of teleology to define the appropriateness of an explanation.

It is argued here that a functional-teleological explanation about the purpose of an organism's trait can be helpful, as a trait can serve a function for the organism. For example, it is appropriate to say that seagulls' wings are for flying. Thus, the distinction between appropriate and inappropriate teleological explanations needs to be made at a deeper level than this simple, functional statement. The distinction instead concerns the rationale for the appearance of organisms' traits. An evolutionary explanation for the appearance of a trait, that there was a variation of a trait in the ancestral population (due to genetic variation), is appropriate. However, a design-teleological account for the appearance of a trait that focuses on an organism's intention, it originated due to some form of *want-based behaviour*, would be inappropriate. Want-based (also termed need-based) explanations are not the only form of design teleology, Kampourakis (2020) also identifies intention-based design teleology where the existence of an organism's trait is due to the intention of an external agent (e.g., a supernatural entity). While such explanations are also inappropriate in evolutionary biology, this paper's focus is want-based design-teleological explanations. For both evolutionary and want-based explanations, a related function-teleological explanation can be used (that the trait serves a purpose for the organism) and, for both, the consequent prevalence of a trait in the current population can be explained due to the trait conferring a selective advantage.

Thus, in this study, functional-teleology in biology is treated as appropriate – organisms' traits can be discussed as having a function; they may serve a purpose. What Trommler and Hammann (2020) referred to as epistemological-teleology and used as an explanatory mechanism in biology. When accounting for the emergence of a trait, it cannot be suggested that a trait appeared in the ancestral population due to want-based behaviour (a specific form of design-teleology), that the trait was designed, and created, by the organism to serve a specific purpose. For example, a tiger wanted to grow stripes to use them as camouflage to better catch prey.

The intervention is designed to encourage children to disregard want-based explanations, and inappropriate type of teleological explanation, and understand the appropriateness of an evolutionary explanation for a trait's appearance. Such want-based accounts are a form of design-teleology, along with intention-based accounts (Kampourakis, 2020), and challenging children's endorsement of them in an important step in reducing teleological reasoning in biology.

Method

Design

Study 2 investigated whether discursive activities about the appropriateness of scientific explanations develop children's understanding of appropriate (evolutionary) and inappropriate (want-based) explanations for the existence of organisms' traits. As in Study 1, a 2x2 mixed experimental design was used, with the same within- and between-groups variables. A no-treatment control was used due to the lack of comparable treatment. The minimum sample size was determined by *a priori* power analysis, based on data in Study 1.

Participants

Twenty-four children, aged 8- to 9-years-old (109 to 118 months-old, *M* = 114 months-old, *SD* = 7.65), took part in the study (Female = 13) from a single-form entry primary school in a large town in the East Midlands, UK. Children were randomly assigned to the treatment (n = 12, Female = 6) or no-treatment control condition (n = 12, Female = 7). Participants did not take part in Study 1.

Treatment

Treatment condition participants were randomised into different groups (three or four members) for the intervention sessions to reduce the effect of group. Treatment sessions were based on the generic intervention structure outlined in Table 1, with two sessions covering elephants' trunks and rabbits' ears. Similar to Study 1, sessions included an introduction, a discussion about the Concept Cartoon statements (see Appendix 2) and a plenary where children were asked to evaluate additional statements as good or bad. Talk moves were used to prompt children during the discussion.

Assessment

The procedure was similar to Study 1, at pre-test assessment participants were asked about how different animals have changed over time and if the conversation could be recorded. Participants were told:

> *That animal's ancestors, those that lived a very, very long time ago, like the great-great-great-great-grandparents of modern animals, used to look rather different. For example, giraffes ancestors' used to have much shorter*

necks, but over time they changed and now modern giraffes have long necks. What I have here are a number of cartoons showing explanations for how species have changed, and I just want to you tell me if you think they are good or bad explanations for how species have changed over time.

All participants were assessed using eight explanations (four want-based and four evolutionary) about the existence of two organisms' traits (reindeer' antlers and giraffes' necks), which were not covered in the intervention. Explanation order for a single organisms' trait was randomised for each participant, as was the order of the two organisms. The wording of these statements is covered below. Participants were asked to judge the explanation as good (appropriate) or bad (inappropriate) and provide a rationale for their judgement. At post-test, children were reminded about the pre-test where they had thought about good and bad explanations regarding species change over time.

Materials

Organisms' traits for the assessment and intervention concept cartoons were selected from children's books and literature on children and teleology. Two functions were selected for each trait. For example, *giraffes' necks* had the trait function of 1) for reaching leaves at the tops of trees and 2) for watching out for danger. Appendix 2 details the set of traits and their functions. Set items were randomly allocated to the assessment, intervention discussion (concept cartoons) or intervention plenary activities. The actual explanation as to why organism X has trait Y differed depending upon whether the statement was a want-based or an evolutionary account.

The want-based accounts for the appearance of traits, those based on design-teleology, stated that a species' ancestors *decided* to develop a specific trait because they *believed* it to be beneficial. Evolutionary accounts for the appearance of a trait focused on there is natural variation in how pronounced a trait was in the ancestral population. Therefore, the first part of each statement detailed the appearance of the trait in the ancestral population. Following this, in both want-based and evolutionary accounts, natural selection was used as the mechanism for the prevalence of the trait in the modern population – that the pronounced version of the trait conferred a selective-advantage. This is a scaled-down version of natural selection considered appropriate for the age-range (e.g., focusing on the adaptation of one trait to an environment across a limited period) and did not include multiple aspects of evolution (e.g., geographic isolation). Using a scaled-down but directed and precise version of natural selection is advocated by Kampourakis and Zogza (2009) and Kelemen et al. (2014). Table 4 details the wording of statements.

Table 4

Study 2 Concept Cartoon statement phrasing

	Want-based	Evolutionary
Abstracted	X's ancestors were all born without Y. However, some X's thought that Y would be good for P. So they started to grow Y to do P. The Xs with Y did P more and so survived to have more babies. Their babies also had Y. Over a long time, all Xs were born with Y, which are for P.	When X's ancestors were born, there was variation in Y. Some Xs had less pronounced Y, some had more pronounced Y, but most were in the middle. Those with pronounced Y were better at P. The Xs with Y did P more and so survived to have more babies. Some of their babies also had Y. Over a long time, all Xs were born with Y, which are for P.
Example	Camels' ancestors were all born without humps. However, some camels thought that humps would be good for storing fat in. So over a number of years they all started to develop humps for storing up fat for when there was no food. The camels with humps stored more fat and so survived to have more babies. Their babies also had humps. Over a long time, all camels were born with humps, which are for storing fat for when there's no food.	When camels' ancestors were born, there was variation in hump size. Some camels had small humps, some had larger humps, but most were in the middle. Those with large humps were better at storing fat for when there was no food. The camels with large humps stored more fat and so survived to have more babies. Some of their babies also had large humps. Over a long time, all camels were born with humps, which are for storing up fat for when there's no food.

Measures

Judgements were coded as in Study 1 (i.e., want-based explanations should be judged as inappropriate and evolutionary explanations as appropriate). Rationales were coded using the rubric in Table 5. Second blind coding of 10% of the data (40 responses) indicated a high level of agreement (K = .931, p < .001).

Table 5

Study 2 rubric for rationales

Score	Explanation type	Example
1	The rationale for the judgment of a statement discusses the style of explanation used in the statement. An evolutionary explanation is good because it offers a causal explanation of how the trait developed – through variation in the population and natural selection. A want-based explanation is bad because it does not offer a causal explanation for how the trait developed. Instead, it states that some of the species wanted the trait as it was beneficial	[Evolutionary explanation for giraffes necks] is good cos says there is variation in the past, and those with long necks lived long and the short necks ones died out [Want-based explanation for reindeer antlers] is bad because it is just saying that they wanted bigger antlers and just decided to start growing them
0	Any other form of rationale including but not limited to a non-response or focusing solely on trait function.	[Evolutionary explanation for giraffes necks] is good because it means that the giraffes can get more food to eat [Want-based explanation for reindeer antlers] is good cos that is a fact, everyone knows that they fight with them

Procedure

Data collection took place across five weeks (June to July). The timetable for the treatment group was the same in both schools:

- Week one, pre-test – individual assessment;
- Week two, treatment session 1 – elephants' trunks;
- Week three, break – school half-term;

- Week four, treatment session 2 – rabbits' ears;
- Week five, post-test – individual assessment.

The no-treatment control group took the pre- and post-tests. For assessment and treatment session, children were taken out of their classroom while the rest of the class continued with lessons unrelated to the intervention. Pre- and post-test individual assessments lasted approximately seven minutes, and group treatment sessions lasted approximately 15 minutes. All school activities were conducted by one the researchers, who has a background as an Early Years teacher.

Results

Two variables were created for the following analyses: judgement score and rationale score (range 0-8). These were the participants' mean score for their eight judgements or eight rationales.

Children's judgements

A 2x2 mixed ANCOVA was conducted to investigate the influence of the intervention on children's judgements about the existence of organisms' traits. The within-groups independent variable was time (pre-test and post-test), the between-groups independent variable was condition (control or treatment). The covariate was children's age in months, which was mean centred to avoid compromising the main effect. The dependent variable was the participants' judgement score.

The analysis revealed a main effect of time (pre-test vs. post-test), F (1, 21) = 48.63, p. < .001, η_p^2 = .698, condition (control vs. treatment), F (1, 21) = 14.30, p. = .007, η_p^2 = .296, and a significant interaction between time and condition, F (1, 21) = 43.52, p. < .001, η_p^2 = .615, explored further with Bonferroni corrected post hoc tests. There was no difference between pre-test scores, *t* (22) = 1.75, *p*. > .999, *d* = .71, but a significant difference between post-test scores, *t* (22) = 5.92, *p*. < .001, *d* = 2.42. Paired *t*-tests showed no significant difference in children's correct endorsement of explanations for the control condition scores from pre- to post-test, *t* (11) = .60, *p*. > .999, d_{rm} = 0.18; however, there was a significant difference in the treatment condition scores, *t* (11) = 12.93, *p*. < .001, d_{rm} = 3.98. There was no significant effect of the covariate (mean centred age in months, evaluated at -.08), *F* (1, 21) = .30, *p*. = .589, η_p^2 = .014, nor a significant interaction between time and age, *F* (1, 21) = 0.13, *p*. = .714, η_p^2 = .007. Table 6 shows the mean judgment scores by condition.

Table 6

Study 2 mean judgement and rationale scores, by time and condition (Max. score = 8)

	Mean judgment score	
	Control	Treatment
Pre-test score (*SD*)	4.05 (1.07)	3.19 (2.91)
Post-test score (*SD*)	4.34 (1.35)	7.48 (1.35)
	Mean rationale score	
	Control	Treatment
Pre-test score (*SD*)	0.45 (0.45)	0.13 (0.55)
Post-test score (*SD*)	0.90 (1.52)	5.01 (1.52)

Children's rationales

A 2x2 mixed ANCOVA investigated the influence of the intervention on children's understanding of the elements of appropriate, or inappropriate, explanations. Independent variables remained the same, and the dependent variable was the participants' rationale score.

The analysis revealed a main effect of time (pre-test vs. post-test), *F* (1, 21) = 85.33, *p*. < .001, η_p^2 = .803, condition (control vs. treatment), *F* (1, 21) = 25.64, *p*. < .001, η_p^2 =

.550, and a significant interaction between time and condition, F (1, 21) = 53.53, $p. < .001$, $\eta_p^2 = .718$, explored further with Bonferroni corrected post hoc tests. Independent t-tests showed no difference between the pre-test scores, t (22) = -1.11, $p. > .999$, d = .46, but a significant difference between post-test scores, t (22) = 1.75, $p. < .001$, d = 2.83. Paired t-tests showed no change in the control condition scores, t (11) = 1.14, $p. > .999$, d_{rm} = 0.18; however, there was a significant change in the treatment condition scores, t (11) = 13.21, $p. < .001$, d_{rm} = 3.98. That is, they provide a higher number of rationales that demonstrated awareness of appropriate or inappropriate explanations for organisms' traits. There was no significant effect of the covariate (mean centred age in months), F (1, 21) = .03, $p. = .856$, $\eta_p^2 = .002$, nor a significant interaction between time and age, F (1, 21) = 0.27, $p. = .603$, $\eta_p^2 = .013$. Table 6 shows the mean rationale scores by condition.

Examples of correct rationales about evolutionary explanations included (see Appendix 2 for the wording of statements):

- *Giraffes evolutionary one* is good 'cos they survived if they had long necks';
- *Giraffes evolutionary two* is good 'cos like small giraffes, they can't see much so it's hard to survive for them, but not the long necks';
- *Reindeers evolutionary one* is good because 'like there is a variation in size, so you can just look at the size and they can just fight, the ones with the big antlers will fight but the small ones will have a bit of a disadvantage';
- *Reindeers evolutionary two* is good because 'I just like that again cos it's not saying they just started to grow but that there was variation'.

Examples of correct rationales about want-based accounts included:

- *Giraffes want-based one* is bad because 'it's just saying they decide to grow it.

You need to put how, how, how do they grow it, how. Cos I want wings, I want a money tree as well coming out my butt, hmmm, see, see it's not happening, where are my wings, where's my money tree?'

- *Giraffes want-based two* is bad because 'they can't just start to grow long necks'.
- *Reindeers want-based one* is bad because 'for fitting with enemies is good, cos I've seen that, but cos they just started to grow antlers, it doesn't make any sense'.
- *Reindeers want-based two* is bad because 'nothing can grow really, we can grow but antlers on reindeers, they wouldn't just go let's grow antlers and they would pop out'.

For comparison, the following are examples of incorrect rationales.

- Some supported the desire of organisms to change: *Giraffe want-based two* is good because "they made their necks grow even longer so they can eat their food" or *Reindeers want-based one* is good "because they say that they are born with small antlers but they made them get bigger".
- Others focused on the function of the trait rather than the mechanism of change:
- *Giraffe evolutionary two* is good "cos it's something most people know already cos if you see one you think that's probably why, so they can reach up."
- Finally, some rationales were unrelated: *Reindeers evolution one* is good "cos there is some good describing words, like instead of big it uses the word large."

Study 2 summary

The results showed that the intervention was successful. Children in the treatment group showed a considerable increase in their correct judgements about the development of organisms' traits. Their rationales also showed a considerable increase in their understanding of the elements of appropriate, and inappropriate, explanations about organisms' traits. Correct rationales for want-based explanations showed that children can recognise that traits cannot come into existence because an organism desires that trait to perform a particular function.

Regarding correct rationales about evolutionary explanations, some children recognised that variation within the ancestral population was integral to the origin of a trait (e.g., see *reindeers evolutionary one & two,* above). Other children focused on the functional advantage that a trait provides (e.g., see *giraffe evolutionary one & two,* above). In the context of scientific accounts for the evolution of traits, both of these elements are vital (other elements such as genetic mutation were not covered by the Concept Cartoon statements). While rationales focusing on the functional advantage of a trait do not reference the mechanism of trait origin (variation), children were still correctly endorsing a scientific account for evolution and identifying that trait function can provide an advantage. This is suggested to represent the start of understanding scientific accounts for trait evolution. The intervention outlined here was not intended to exist in a vacuum; further educational activities would be required to support children in developing their understanding of evolution. This intervention can, however, act as a first step in recognising correct elements of scientific explanation.

The inferences from Study 1 about the intervention design hold. At its core, the intervention was simple; it required neither costly resources nor significant teaching time. After only 30 - 40 minutes of dialogic activity, children showed considerable improvements in their understanding of explanations for the existence of organisms' traits.

Discussion

Both interventions were successful in limiting children's endorsement of and developing their understanding of teleological explanations. In the two treatment conditions, children showed significant improvement from pre-test and considerably outperformed those in the control condition at post-test. These significant differences were supported by very large effect sizes. The results hold for children's initial judgements about the 'correctness' of different explanatory styles and their understanding of why explanations are appropriate or inappropriate.

Children in the treatment conditions approximately doubled the number of correct judgements they provided (44% to 83%, in Study 1, 40% to 94%, in Study 2). This considerable improvement was matched in the number of correct rationales that children provided to support their judgements (17% to 64%, in Study 1, 2% to 63%, in Study 2). In Study 1, the intervention had a dramatic impact on children's understanding, with over two-thirds of children explaining why teleological explanations should not, and why scientific explanation should, be used to explain the existence of natural phenomena. In Study 2, before the intervention children did not understand what constitutes appropriate explanation about the emergence of organisms' traits. However, following two short sessions, two-thirds of children could explain why want-based explanations are inappropriate and why evolutionary explanations are appropriate.

Existing work on teleology typically catalogues children's use and endorsement of teleological explanations (e.g., see Kelemen, 1999) without attempting to change children's thinking. Undoubtedly, such studies are useful for exploring the ideas that children bring to learning environments. However, using the results of such studies to decide what children can learn would be misleading, as this would suggest teleology is a barrier to young children understanding specific theories. The findings of this paper show that with support, children can recognise and disregard some forms of inappropriate teleological explanations from a young age.

Children’s teleological preconceptions were not inherently difficult to change; a finding that runs counter to current thinking on children and teleology. Work on children’s preconceptions related to scientific concepts (of which teleological explanations belong) suggests that preconceptions are highly resistant to change and pose problems when children learn about scientific theories (e.g., see Holman & Yeomans, 2018). This, however, was not the case. It is argued that targeting a preconception and specifically exploring why it is incorrect, rather than trying to teach the correct scientific account, is an effective method to reduce children’s endorsement of teleological explanations. Furthermore, the results of Study 2 show that children’s endorsement of want-based explanations (a form of design-teleology) can be challenged. Children can be taught to recognise want-based statements and criticise their use as scientific explanations. Embedded within a biology curriculum, alongside specific instruction on evolution, this type of intervention could be used to develop an understanding of teleology in biology, as per suggestions by Trommler and Hammann (2020).

The generic intervention was shown to be effective across two different age ranges (5- to 7-year-olds and 8- to 9-year-olds) and for two different categories of topic (natural phenomena and organisms’ traits). Consequently, the generic intervention structure and method could be adapted for other age-ranges; indeed, there are studies which demonstrate that high school and undergraduate students can be taught to distinguish between teleological and causal explanations (see Bartov, 1978; Richardson, 1990).

Limitations

The results presented here show the efficacy of the interventions only in the respect that they were shown to work better than a no-treatment control and hence demonstrate the findings are not merely due to repeated testing or maturation. Consequently, future work is needed to test relative efficacy. The activities used in the interventions may be amongst the most effective activities or (less plausibly given the effect size) may be amongst the least effective. Future studies are required which compare the intervention presented here with active control conditions or combine it with comparable interventions that seek to develop similar conceptual knowledge. These may directly teach children the correct scientific explanations (e.g., Kampourakis & Zogza, 2009; Kelemen et al. 2014). Work is also required to explore how teachers adapt the materials to their classroom. For example, how concept cartoon statements can be tailored to the preconceptions of their pupils.

As noted, the focus in Study 2 was want-based explanations, which only represent one form of design-teleology. To effectively tackle the problem of inappropriate teleological reasoning in evolutionary biology, other forms of teleological explanation must be challenged. Future work could explore how the intervention outlined here or other teaching methods (e.g., Kelemen et al. 2014) could challenge children's endorsement of intention-based design teleology (see Kampourakis, 2020). Finally, these studies focused on children's endorsement of teleological explanations. However, children do not only evaluate explanations but also generate their own. An important next step is evaluating how the intervention influences children's generation of teleological explanations.

Conclusion

The studies reported here show that dialogic activities can improve children's understanding of explanatory styles in science. Across two studies, children took part in dialogic activities that developed their understanding of why inappropriate forms of teleological explanation should not be used in scientific contexts. Running counter to suggestions in existing literature, children's endorsement of inappropriate forms of teleological explanations was decreased after a series of short, simple discursive interventions. Educators can use such discursive activities to reframe teleology and challenge this pervasive preconception.

Acknowledgements

This work was supported by the Economic and Social Research Council [grant number 1366484].

Declaration of interest statement

The authors have no potential conflict of interest associated with this work.

References

Alexander, R. (2006). *Towards dialogical teaching: Rethinking classroom talk* (3rd ed.). York: Dialogos.

Barnes, M.E., Evans, E.M., Hazel, A., Brownell, E.R. & Nesse, R. (2017). Teleological reasoning, not acceptance of evolution, impacts students' ability to learn natural selection. *Evolution: Education and Outreach, 10*(7), 2-12. doi:10.1186/s12052-017-0070-6

Bartov, H. (1978). Can students be taught to distinguish between teleological and causal explanations? *Journal of Research in Science Teaching, 15*(6), 567-572.

Chin, C. & Teou, L.Y. (2009). Using concept cartoons in formative assessment: Scaffolding students' argumentation. *International Journal of Science Education, 31*(10), 1307-1332. https://doi.org/10.1080/09500690801953179

diSessa, A.A. (2014). A history of conceptual change research: Threads and fault lines. In R.K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (2nd ed., pp. 88-108). Cambridge: Cambridge University Press

DiYanni, C. & Kelemen, D. (2005). Time to get a new mountain? The role of function in children's conceptions of natural kinds. *Cognition, 97*(3), 327-335. doi:10.1016/j.cognition.2004.10.002

Emmons, N., Smith, H. & Kelemen, D. (2016) Changing Minds With the Story of Adaptation: Strategies for Teaching Young Children About Natural Selection, Early Education and Development, 27(8), 1205-1221. doi: 10.1080/10409289.2016.1169823

Galli, L.M.G. & Meinardi, E.N. (2011). The role of teleological thinking in learning the Darwinian model of evolution. *Evolution: Education & Outreach, 4*, 145-152. https://doi.org/10.1007/s12052-010-0272-7

Grugeon, E., Hubbard, L., Smith, C. & Dawes, L. (2012). *Teaching speaking and listening in the primary school*. Routledge: Oxon.

Halls, J.G., Ainsworth, S.A. & Oliver, M.C. (2018) Young children's impressionable use of teleology: the influence of question wording and questioned topic on teleological explanations for natural phenomena. *International Journal of Science Education*, 40(7), 808-826. https://doi.org/10.1080/09500693.2018.1451008

Hanke, D. (2004). Teleology: The explanation that bedevils biology. In J. Cornwell (Ed.), *Explanations: Styles of explanation in science*. Oxford: Oxford University Press.

Henderson, J.B., MacPherson, A., Osborne, J. & Wild, A. (2015). Beyond construction: Five arguments for the role and value of critique in learning science. *International Journal of Science Education, 37*(10), 1668-1697. https://doi.org/10.1080/09500693.2015.1043598

Holman, J. & Yeomans, E. (2018) Improving Secondary Science; Guidance report. London: Education Endowment Foundation.

Howe, C. & Abedin, M. (2013). Classroom dialogue: A systematic review across four decades of research. *Cambridge Journal of Education*, 43, 325-356. https://doi.org/10.1080/0305764X.2013.786024

Kampourakis, K. (2014). *Understanding evolution*. Cambridge: Cambridge University Press.

Kampourakis, K. (2020). Students' "teleological misconceptions" in evolution education: why the underlying design stance, not teleology per se, is the problem. Evolution: Education and Outreach, *13*(1), 1-12. https://doi.org/10.1186/s12052-019-0116-z

Kampourakis, K., Pavlidi, V., Papadopoulou, M. & Palaiokrassa, E. (2012). Children's teleological intuitions: What kind of explanations do 7–8 year olds give for the features of organisms, artifacts and natural objects? *Research in Science Education, 42*(4), 651-671. doi:10.1007/s11165-011-9219-4

Kampourakis, K. & Zogza, V. (2009). Preliminary evolutionary explanations: A basic framework for conceptual change and explanatory coherence in evolution. *Science & Education, 18*(10), 1313-1340. doi:10.1007/s11191-008-9171-5

Kelemen, D. (1999) Function, goals and intention: children's teleological reasoning about objects. *Trends in Cognitive Sciences, 3*(12), 461-468. 10.1016/s1364-6613(99)01402-3

Kelemen, D., Emmons, N.A., Seston Schillaci, R. & Ganea, P. A. (2014). Young children can be taught basic natural selection using a picture-storybook intervention. *Psychological Science, 25*(4), 893-902. doi:10.1177/0956797613516009

Loxley, P., Dawes, L., Nicholls, L. & Dore, B. (2010). *Teaching primary science: Promoting enjoyment and developing understanding*. Harlow: Pearson Education Limited.

Mercer, N., Dawes, L., Wegerif, R. & Sams, C. (2004). Reasoning as a scientist: Ways of helping children to use language to learn science. *British Educational Research Journal, 30*(3), 359-377. doi:10.1080/01411920410001689689

Mercer, N. & Howe, C. (2012). Explaining the dialogic processes of teaching and learning: The value and potential of sociocultural theory. *Learning, Culture and Social Interaction, 1*(1), 12-21. https://doi.org/10.1016/j.lcsi.2012.03.001

Michaels, S. & O'Connor, C. (2012) *Talk Science Primer.* Cambridge: TERC.

Naylor, S. & Keogh, B. (1999). Concept cartoons, teaching and learning in science: An evaluation. *International Journal of Science Education, 21*(4), 431-446. https://doi.org/10.1080/095006999290642

Naylor, S. & Keogh, B. (2000). *Concept cartoons in science education*. Sandbach: Millgate House Publishers.

ojalehto, b., Waxman, S.R. & Medin, D.L. (2013). Teleological reasoning about nature: Intentional design or relational perspectives? *Trends in Cognitive Sciences*, 17(4), 166-171.

Richardson, D. R. (1990) A survey of students' notions of body function as teleological or mechanistic. *Americal Journal of Physiology, 258*, 8-10.

Ruse, M. (1989) Teleology in biology: Is it a cause for concern? *TREE, 4*(2), 51-54. https://doi.org/10.1016/0169-5347(89)90143-2

Trommler, F. & Hammann, M. (2020) The relationship between biological function and teleology: implications for biology education. Evolution: Education and Outreach, 13(11) https://doi.org/10.1186/s12052-020-00122-y

Trommler, F., Gresch, H. & Hammann, M. (2017) Students' reasons for preferring teleological explanations. International Journal of Science Education, 40(2), 159-187. https://doi.org/10.1080/09500693.2017.1404658

Zohar, A. & Ginossar, S. (1998). Lifting the taboo regarding teleological and anthropomorphism in biology education - heretical suggestions. *Science Education, 82*, 679-697. https://doi.org/10.1002/(SICI)1098-237X(199811)82:6<679::AID-SCE3>3.0. CO;2-E