

We're going to the zoo: Interactive educational activities with animals boost category-based induction in children

Julia R. Badger*

Department of Experimental Psychology, University of Oxford, Oxford, OX1 3UD, UK

*Corresponding Author: julia.badger@psy.ox.ac.uk; +44 (0) 1865 271417

Laura R. Shapiro

School of Life and Health Sciences, Aston University, Birmingham, B4 7ET, UK

l.r.shapiro@aston.ac.uk; +44 (0) 121 204 4052

The research took place at Aston University and was funded by an Economic and Social Research Council Studentship, and a Twycross Zoo Conservation Fund, awarded to the first author. We would like to thank the Education Team at Twycross Zoo and the staff and pupils at our participating schools for their contributions to this work.

Word count (exc. figures/tables): 5,598

Abstract

We investigated whether category-based induction can be enhanced through educational activities with real-life animals. Four induction tasks involving pictures of real and novel biological kinds were administered to 252 children aged 5- to 7- years, split across two testing sessions. Between these two testing sessions, 129 of these children took part in a zoo-based educational activities where their attention was directed towards the importance of non-obvious category features. In the first testing session, older children made significantly more category inferences, consistent with developmental accounts of category-based induction. In the second testing session, there was a greater increase in category-based induction decisions made by children in the training condition. We suggest that category-based induction could be driven by an accumulation of category knowledge which leads to a deeper understanding of the importance of category features.

Keywords: Induction; Education; Training; Child development; Category

Introduction

Inductive reasoning is a fundamental cognitive skill that enables people to make sense of the world by applying their existing knowledge to novel situations. Imagine finding a canned cheeseburger in your local supermarket. This may be entirely new to you, but from your knowledge of other canned foods, you'll induce that it can be kept in a cupboard, not the fridge, and if left unopened it will probably last a long time. Our ability to make these inferences is called category-based induction and is an essential part of much of our everyday reasoning (Hayes, Heit & Swendsen, 2010). A large body of research focuses on category-based induction in children, yet there remains considerable disagreement about how this ability develops. Some proponents of the naïve (folk) theory account argue that early-learned biases lead even very young children to show an early preference for category-based induction for natural kinds (Bulloch & Opfer, 2009; Gelman, 2003; Gelman & Markman, 1986). In contrast, proponents of the similarity account argue that a category-based induction preference develops gradually (Badger & Shapiro, 2012, 2015; Sloutsky, Kloos & Fisher, 2007; Sloutsky, Deng, Fisher & Kloos; 2015). The similarity account draws on Piaget's proposed shift from preoperational to concrete operational thinking (Inhelder & Piaget, 1958; Piaget, 1964), and emphasises that children are similarity-focused when young and will therefore make inferences based on the overall similarity of the items presented (whether that be perceptual or representational similarity; Fisher, Godwin & Matlen, 2015). Even when the categories presented are familiar, if there is a perceptually or linguistically similar match available, young children will make similarity-based inferences and only make categorical inferences under less demanding conditions (Godwin & Fisher, 2015; Long, Zhang & Deák, 2012; Sloutsky & Fisher, 2004). According to this account, children's use of category-level information for induction decisions increases with development (Badger and Shapiro, 2012; 2015).

The current study aimed to contribute to this literature in two ways: firstly, by examining children's induction performance for real and novel examples of plausible natural kind categories in which between-category similarity was high, and different categories were distinguished by non-obvious features (e.g., chinchilla versus squirrel). Secondly, by examining whether teaching children to focus on category features can influence the extent to which they make category-based induction

decisions. Below, we discuss research investigating how a child's experience influences their development of category-based induction, raising the possibility that category-based induction may be influenced by teaching.

The Effect of Direct Experience and Training with the Biological World

Previous research has already gone some way to understanding the influence of a child's environment and experience on their use of category-based induction. According to Carey's conceptual change account (Carey, 1985), young children see humans as distinct from other animals (anthropocentrism) and do not develop a cognitive model whereby humans are seen as animals until age 10-12. More recent research supports the view that a child's cultural and environmental experience influences their development of biological knowledge. For example, in 1990, Inagaki found that young children who had experienced looking after a goldfish for a prolonged period of time at home were more likely to make biologically plausible inferences than those who had no experience looking after an animal. This finding was supported by the work of Atran, Medin, Lynch, Vapnarsky, Ucan & Sousa (2001) who found gender differences in children's induction decisions about a peccary (a pig-like mammal). Boys outperformed girls, which the authors hypothesised to be caused by cultural differences in girls' and boys' familiarity with the animals (from a young age, boys spend time in the forest with their fathers where they encounter peccaries more often). Similarly, Ross, Coley, Medin and Atran (2003) worked with a large sample of children from three distinct populations – rural Native American, rural American and urban American – and examined their folk biological induction of unfamiliar properties to pictures of biological kinds (plants, animals, humans). They found that only urban children showed evidence of early anthropocentrism; younger rural children showed evidence of biological affinity, and Native American and older rural children showed evidence of full ecological reasoning. These differences were hypothesised to reflect cultural differences, and the extent to which children's upbringing increased their experience with the natural environment (see also Medin & Atran, 2004). More recently, Coley (2012) found that older children, like adults, were more likely to use taxonomic, categorical information when reasoning about the biological insides of one species compared to another but that they were more likely to use ecological or physical surrounding information when reasoning about diseases spread from one species to another.

However, he found that the emergence of this inductive method occurred earlier among rural children (age 6) than it did in suburban (age 10) or urban children (age 10+). Together, these studies suggest that individual differences in domain-relevant culture and experience can influence the development of biological reasoning. Specifically, the more direct-experience with animals, the more likely children are to use category-induction earlier. This suggests that an accumulation of category knowledge through direct experience increases children's awareness of the importance of category features across the biological domain, driving the development of adult-like category-based induction. However, we cannot make a strong causal argument based on these naturalistic studies because the groups vary in so many respects.

There is also evidence from experimental studies that experience in the form of corrective-feedback can influence children's categorization biases and induction decisions. Bulloch and Opfer (2009) presented 3- to 5-year olds with computer-based triad illustrations of juvenile and adult insects. The juveniles were either dark bodied and pale headed, or pale bodied and dark headed. The adults comprised six segments, each of which could be altered in five ways to allow for the level of perceptual similarity to be adjusted. These were presented alongside two additional triads to create one trial of a target triad and two test triads. Children were given either offspring or prey scenarios to explain the relationship between the adults and juvenile in the test triads and were asked questions regarding categorization, growth projection and hidden property induction of the target juvenile in relation to the test juveniles. They found that 3-year olds showed no difference in their answering for either scenario whereas 4- and 5-year olds had started to understand that the parent-offspring relationship could inform categorization and induction. However, after a feedback phase (completing the same induction task, but with corrective feedback), Bulloch & Opfer found that the 3-year olds' answers were in line with the older children and adults: relational matching for the parent-offspring scenario but not for the hunter-prey condition. More recently, Lawson, Fisher and Rakison (2015) found that they could influence toddlers' categorization biases by teaching them to group items using observable features (e.g., has a trunk) or unobservable features (e.g., is clever). In the pre- and post-tests children were shown a target item and asked to match it to one of two test items (described in terms of both observable and unobservable features). Most children at pre-test showed no bias or a

perceptual bias, but at post-test, children's biases were consistent with their training condition.

Although this study investigated categorization rather than induction decisions, it showed that very young children can be trained to focus on unobservable features. However, these studies specifically trained children on the task, in an experimental situation. Together, the naturalistic and the experimental studies described above suggest that children can, even with limited exposure, learn to generalize on the basis of unobservable relations. We build on this work and examine the extent to which training on natural kind categories can influence children's category-based induction even when tested outside the context of the training situation.

Current Study

In a quasi-experimental study using novel biologically plausible categories taken directly from Badger and Shapiro (2012; 2015) alongside stimuli of real biological kinds, we aimed to investigate whether category-based induction can be influenced by a one-off direct training experience with animals. Specifically, we explored whether interactive educational activities at the zoo, explaining how animals have adapted to their habitats by focusing on key features and behaviors, increased 5- to -7-year-old children's use of category-based induction in the biological domain: an experimental study based on a short naturalistic experience.

Following on from research showing that a) category-based induction develops with age (Badger & Shapiro, 2012), and that b) prolonged exposure to the biological world can directly influence biological reasoning (Coley, 2012; Inagaki, 1990; Ross et al., 2003), we predicted that there should be room for improvement in young children's category-based induction. We investigated the influence of a 30-minute zoo-based educational activity which aimed to teach children to focus on non-obvious category features, which have importance for animal behavior, and the way in which they have adapted to their habitats. There are two potential ways in which such an activity may cause category-based induction to increase. First, the activity may increase children's awareness of the importance of category features and encourage them to focus on these features when making induction decisions (general awareness). This would lead to a general increase in category-based induction post-training for tasks using both real kind and novel kind stimuli (examples of real animals

from the training, or novel untrained categories). Second, the training may increase children's awareness of the importance of category features, but only for the categories on which they have been trained (specific awareness). They would then be more likely to focus on these features when making induction decisions that involve these trained categories, showing an increase in category-based induction post-training, but only for the specific categories included in the educational activities. This would not tell us how children ultimately generalize their category-based induction preferences to unfamiliar categories. Nevertheless, one plausible route to a general category-based induction bias would be via a gradual accumulation of specific knowledge, which would ultimately generalize to a category-based induction bias across the biological domain.

Triad paradigms are commonly used to measure category-based induction, in which the child is shown a target and two test items, one of which matches the target by category but not perceptually, and one of which matches the target perceptually but not by category. For example, a target blackbird could be shown alongside a dodo (same category as the target, but perceptually dissimilar) and a bat (different category, but a closer perceptual match). Children must generalize a hidden property from the target to one of the test items. If the child matches the target to the category test item, their inference is categorical. If the child matches the target to the perceptual test item, their inference is perceptual. The methodology of this study followed very closely those used by Badger and Shapiro (2012; 2015) and Sloutsky, Kloos and Fisher (2007) whereby the child completes four phases for each induction task: 1) the category learning phase, 2) the initial categorization phase, 3) the induction phase which uses the triad paradigm, 4) the final categorization phase (see Procedure for more detail).

Method

Participants

Previous research studies using the same broad design reported that the transition from perceptual to category-based induction occurs between ages 5 and 8 (Badger & Shapiro, 2012; 2015). The interactive educational activities at Twycross Zoo are targeted at children in U.K. school years 1 and 2, spanning a similar age range. A total of 252 children from six English primary schools

participated in our quasi-experimental design: 126 year 1 (5;7-6;10 years; M = 6;3) and 126 year 2 (6;2-7;9 years; M = 7;2): 124 males. Of these, 123 participated in the comparison condition and 129 in the training condition, providing power of >95% to detect medium effect sizes. The training condition took part in an interactive educational activity at Twycross Zoo, and both conditions completed four counterbalanced tasks, two pre- and two post-training period (session 1 and session 2). The training group was recruited through Twycross Zoo: the children's schools had already booked their year 1 and year 2 classes to attend the interactive educational activities at the Zoo. Each activity was delivered to a whole class of children, within their usual school classes. The comparison group was recruited separately from nearby schools matched in terms of their age, location, school type, National Curriculum achievement in English and Mathematics and the school's Ofsted report (an inspection report generated by the Office for Standards in Education, Children's Services and Skills, on a school's overall effectiveness).

Stimuli

Four sets of stimuli were used, each including two contrasting categories. Two sets of novel items never seen by the children before (sandbug vs. rockbug, taken from Badger & Shapiro, 2012; ground weevil vs. forest weevil, taken and renamed from Badger & Shapiro, 2015) and two sets of real kind items (chinchilla vs. squirrel; tree frog vs. common frog). Both real kind item sets used photographs of real animals, pasted onto a white background. Using both real and novel kind stimuli with the same children enabled us to directly test whether realistic computer aided design (CAD) novel kinds result in the same use of category-based induction as real kinds or whether they are seen, and therefore treated, differently. If category-based induction is a bias, then children should be able to use it regardless of whether an item is a known real-life category or a biologically plausible newly learned category.

[Figure 1 here]

These sets of stimuli were used to create four induction tasks, following Badger and Shapiro's (2012; 2015) procedure. In each induction trial, the child was first shown an animation in which a juvenile animal transforms into an adult animal (Figure 1), followed by a triad including an

adult target, an adult test item (perceptual distractor) and a juvenile test item (category inference; see Figure 2). It has been shown that children are not distracted by the age-category of the stimuli in these triads (juvenile vs. adult test choice), and instead it provides a neat way to separate perceptual versus categorical inferences (Badger & Shapiro, 2012).

[Figure 2 here]

To check that the perceptual distractor was seen as perceptually more similar to the target than the category choice, we conducted a similarity pre-test in which 31 children not taking part in the main study (11 reception, 10 year 1 and 10 year 2; range 4;2-7;1; 15 males and 16 females) were shown all 48 triads and asked which test item looked most like the target in each. One-sample *t* tests confirmed that in all triads the perceptual choice was chosen significantly more than chance, validating them as perceptual distractors, see Table 1. This was true for all the year groups. It is important to note that the perceptual choice was chosen more often for the novel than the real kinds, suggesting that there was a stronger perceptual lure for these conditions. This finding precludes any direct comparisons between real and novel items, per se. However, we can still examine whether we observe any age-effects or training effects for each type of stimuli. The same children were then shown 32 images individually (two juveniles and two adults from each category) and were found to assign biological names to every image, both real and novel. Real kinds were identified as 'frogs', 'squirrels', 'chinchillas' and 'rabbits'. Novel kinds were identified as 'bugs', 'beetles', 'spiders' and 'worms'.

[Table 1 here]

Procedure

The 129 children in the training condition took part in a real interactive educational activity at Twycross Zoo. Children in both the comparison and the training conditions completed all four induction tasks: two a week before (one real and one novel; session 1) and two a week after the educational zoo activity (one real and one novel; session 2). Tasks were counterbalanced to avoid order effects: children either completed the common frog-tree frog task and sandbug-rockbug task during session 1 and then the chinchilla-squirrel and ground weevil-forest weevil tasks during session

2, or vice versa. All 252 children completed the session 1 and session 2 tests in a quiet area outside their classrooms at school. Testing was completed individually with the researcher using an E-Prime-programmed task on a laptop. The testing environment was kept as similar as possible, across schools, children and sessions.

Each induction task began with a demonstration of the two categories (e.g., chinchilla vs. squirrel), and their key features were explained. In line with previous research, children were given categorization rules for each category of stimuli as a method for discrimination between the two categories presented on screen. They saw a juvenile item, for example a chinchilla, appear on screen and were told: “Here is a chinchilla. Chinchillas live in cold mountains and are round and fluffy for keeping warm. Let’s watch it grow up”. Then the child would watch the chinchilla transform from a juvenile into an adult (see Figure 1 for examples of transformations): “Now it has grown up, it is still a chinchilla. It lives in cold mountains and is round and fluffy for keeping warm”. Children were only told the category membership rules for the two categories used in a specific task, immediately prior to that task. These were the rules (category information / knowledge) for all four tasks:

1) Sandbugs live in the sand and have round heads for making soft burrows; rockbugs live rocks and have angled heads for digging.

2) Forest weevils live in trees and have long tails for good balance; ground weevils live in the earth and have short tails to fit into small holes.

3) Tree frogs live in trees and have sticky pads on their feet for climbing; common frogs live in ponds and have straight webbed feet for swimming.

4) Squirrels live in trees and have long bushy tails for balancing; chinchillas live in cold mountains and are round and fluffy for keeping warm.

After this demonstration, the children completed four phases for each induction task: 1) the category learning phase, 2) the initial categorization phase, 3) the induction phase, 4) the final categorization phase. Different stimuli were used during every phase to avoid the effect of recognition.

In the category learning phase, children were given 8 trials (4 of each category), in random order, in which they were asked to name the item's category at the juvenile and adult stages of transformation (see Figure 1 for examples). This provided a check that children understood the continuity between juvenile and adult. Feedback on whether the child's answer was correct or incorrect was given and then they were reminded of the rule after each answer.

In the initial categorization phase, children were given 12 trials in the same format as the category learning phase, (6 of each category), in random order, and were asked to name the category of the juvenile followed by the adult. They were given no feedback and the rules were not reiterated.

In the induction phase, children completed 12 trials (see Figure 2 for examples). A juvenile would transform into an adult (the target). Then the two test items would appear, forming a triad on-screen. When the triad appeared, the child was told a hidden property of the target (see Appendix 1) and asked by the researcher to decide which test item shares this hidden property. No category labels were given. If the child pointed to the perceptual distractor item, this was coded as a perceptual inference; if the child pointed to the category item, this was coded as a categorical inference.

In the final categorization phase, children were given 12 trials following the same format as the initial categorization phase. This provided a check that children could still remember the rules for differentiating between the two categories after the induction phase.

Interactive educational Activities

Children in the training condition received an educational activity called 'Animal Adaptations to Habitats'. The aim of the educational activity was to highlight unique features of each animal and to link these features to functions and habitats. This was a 30-minute activity delivered to a class of approximately 30 children in the educational centre at Twycross Zoo, designed and presented by members of the educational team. Children saw, touched and learnt about five small animals (see Figure 3). The educational team discussed the importance of each animal's features for their environment. The educator asked the children questions and talked about a specific animal whilst pointing to the feature of importance, for example, "a tree frog has sticky pads on its feet to climb trees" [pointed to pads on feet]. They then took the animal to every child to hold, touch or look at,

whilst again pointing out the feature discussed. Every child discussed, and was taught, the information in their class group, and then had the opportunity to re-identify the new information on a one-to-one basis.

These activities were not adapted to our tasks. Instead, we adapted our tasks to include examples of the animals seen during the activities (and a familiar perceptual match for each, e.g., squirrels to perceptually match the chinchillas seen during the activity), and the features that were highlighted during the activities were used as our category membership rules.

[Figure 3 here]

Analysis Strategy

Our testing sessions were delivered to classes of children from six schools. We therefore examined the effect of this school-level clustering before deciding on our analysis approach. A variance components analysis was conducted to examine the proportion of variance explained by each of our fixed and random factors on the percentage of category-based induction decisions. There were two within-subjects factors: test session (with two levels: session 1 and session 2; note that session 2 was post-training for the experimental group, and after an equivalent gap for the comparison group) and kind (with 2 levels: real vs. novel). There were two between-subjects factors: year group (with two levels: year 1 vs. year 2) and condition (with two levels: training vs. comparison). There were two random factors: children and schools. We found that 56% of the variance in category-based induction decisions was explained by individual differences between children. A further 35% of variance was explained by interactions between children and our fixed factors. No variance was explained by individual differences between schools alone, although 9% of variance was explained by interactions between schools and our fixed factors (2% school*session; 3% school* year; 4% school*year*kind*session). Since school-level clustering explained such a small amount of variance, we report single-level analyses, including a repeated-measures GLM to examine the effects of condition on category-based induction decisions.

Results

Categorization Performance

This section describes the two categorization phases, administered at the beginning and end of each testing session. The purpose was to check that children could accurately categorize the items before they started the induction phase, and that they could still do so at the end of the session, after they had completed the induction phase.

According to a binomial test, scores of 10/12 (83%) and above were significantly above chance (proportion = 0.5, $p = .04$). Only children who performed significantly above chance in both initial and final categorization phases in a testing session were included in the final sample. This was to ensure that they had understood (initial) and remembered (final) the categories during the induction phase. Children had to pass the categorization phases for all four tasks to be included in our analyses. Seventeen children (12 Y1 (5 to 6-year-olds), 6 training and 6 comparison; 5 Y2 (6- to 7-year-olds), 2 training and 3 comparison) scored at or below chance during the categorization phases and were removed. The remaining 235 children (114 comparison, 53 year 1, 61 year 2; 121 training, 61 year 1, 60 year 2) scored highly and above chance on the initial and final categorization phases of testing: **Real (first testing session):** Initial $M = 96\%$ ($SD = 5.27$), $t(234) = 133.05$; $p < .001$; Final $M = 97\%$ ($SD = 5.08$), $t(234) = 141.52$; $p < .001$. **Real (second testing session):** Initial $M = 94\%$ ($SD = 6.09$), $t(234) = 111.58$; $p < .001$; Final $M = 95\%$ ($SD = 5.60$), $t(234) = 124.19$; $p < .001$. **Novel (first testing session):** Initial $M = 98\%$ ($SD = 4.02$), $t(243) = 184.09$; $p < .001$; Final $M = 98\%$ ($SD = 4.08$), $t(234) = 181.33$; $p < .001$. **Novel (second testing session):** Initial $M = 98\%$ ($SD = 4.38$), $t(243) = 167.47$; $p < .001$; Final $M = 98\%$ ($SD = 3.91$), $t(234) = 189.16$; $p < .001$.

Induction Performance

This section focuses on children's induction scores for session 1 and session 2 (i.e., pre- and post-training for the training condition), but only for those children who reached criterion on the categorization phases. Our tasks were counterbalanced, enabling us to confirm that there were no effects of task order (two levels: common frog-tree frog and sandbug-rockbug tasks pre-training vs.

chinchilla-squirrel and ground weevil-forest weevil tasks pre-training; $F(1, 233) = 2.66$; $p = .104$, $\eta_p^2 = .011$).

We then conducted a repeated measures GLM to examine the change in real and novel kind induction scores for the comparison and training groups, with specific focus on the interaction between condition and session to see whether the training group made a greater increase in category-based induction. There were two within-subjects factors: session (with two levels: session1, 'pre-training' vs. session2, 'post-training') and kind (with 2 levels: real vs. novel). There were two between-subjects factors: year group (with two levels: year 1 vs. year 2) and condition (with two levels: training vs. comparison). The dependent variable was the percentage of category-based induction decisions.

There was no main effect of condition ($F(1, 231) = 2.02$; $p = .156$, $\eta_p^2 = .009$), indicating that the comparison and training groups showed similar performance overall. There was a significant main effect of testing session ($F(1, 231) = 22.84$; $p = .001$, $\eta_p^2 = .090$), with a greater percentage of category-based induction decisions for the second session, overall. There was a significant main effect of year group ($F(1, 231) = 16.63$; $p = .001$, $\eta_p^2 = .067$), with year 2 showing more category-based induction decisions than year 1 (Figure 4). There was a significant main effect of kind (real vs. novel; $F(1, 231) = 24.20$; $p = .001$, $\eta_p^2 = .095$), with a greater percentage of category-based induction decisions for the real kind tasks, overall.

There was a significant interaction between session and condition ($F(1, 231) = 6.55$; $p = .011$, $\eta_p^2 = .028$) with the training condition showing a greater increase in category-based induction decisions between the two sessions (Figure 4). The interaction between session and year approached significance ($F(1, 231) = 3.26$; $p = .072$, $\eta_p^2 = .014$), indicating a trend towards a greater increase in category inferences for the older year group. All other interactions were non-significant ($F_s < 2.71$; $p_s > .101$, $\eta_p^2 < .012$). Post-hoc t-tests were conducted to explore the significant interaction between session and condition, and the almost significant interaction between session and year. All t-tests were non-significant, correcting for multiple-comparisons ($p > .025$). Thus, the interaction between

session and condition reflects the greater gains made by the training group, but these gains did not translate into an absolute advantage on category-based induction performance.

In sum, the training group showed a significantly greater increase in category-based induction. Children made more category-based induction decisions for the real kinds than the novel kinds but there were no significant interactions with kind, suggesting that the effect of training was stable across real and novel kinds. The older children made more category-based induction decisions than the younger group and there was a trend for the effect of training to be greater for the older group.

[Figure 4]

Discussion

Using zoo-based educational activities, we aimed to explore whether a one-off, short training activity with real biological kinds could enhance the development of category-based induction in the biological domain.

First, we found that category-based induction develops and becomes more robust with age. We observed a shift away from similarity-based towards category-based induction with Year 2 (6- to 7-year-olds) children making more category-based induction decisions than Year 1 (5- to 6-year-olds) children. Therefore, our data are broadly consistent with developmental accounts such as Badger and Shapiro (2012) and Sloutky, Kloos and Fisher (2007). However, these previous studies examined induction decisions for novel kinds, whereas we compared induction decisions for real and novel biological kinds. Although children made more category-based induction decisions for real than novel kinds, this effect did not interact with age, suggesting that children's category-based induction increases for real as well as novel kinds between the ages of 5 to 7.

Second, our main novel contribution was the inclusion of a 30-minute training activity with real animals, highlighting the importance of non-obvious features common to a category and animal behavior. We found a significant interaction between condition and session (small to medium effect size), with the trained group showing a greater increase in the number of category decisions in the

second testing session. Previous research has shown the positive impact of prolonged experience and environment on biological reasoning (Atran et al., 2001; Coley, 2012; Inagaki, 1990; Ross et al., 2003) and that categorization and induction biases can be influenced by feedback during experimental tasks (Bullock & Opfer, 2009; Lawson et al, 2015). We have found the same positive effect from only a short real-world experience in an entirely separate context (direct experience of animals in a zoo) from the follow-up experimental task (computerized task, conducted in a classroom). During the educational activities at the zoo, children held small animals such as chinchillas and tree frogs. They were encouraged to attend to their non-obvious features and the importance of these features for survival in the animal's habitat (conceptual information). Being taught to focus on category features alongside hands-on interaction, appeared to increase the children's use of these features when making category-based induction decisions during post-training induction tasks. These findings are compatible with Fisher et al.'s (2015) theory about the role of representational similarity in inductive reasoning. Specifically, as children learn about the features common to a category, this greater feature overlap leads to greater representational similarity. As their representations of those categories become more finely tuned, this in turn leads to an increased likelihood of category inference. These findings are also compatible with Carey's conceptual change account, whereby children's cognitive model restructures and develops with the accumulation of biological knowledge. In addition, our findings are in line with previous studies emphasizing the importance of real-life experience and culture for category-based induction. We have demonstrated that even a short-term increase in children's experience, coupled with explicit teaching about the importance of category features, can influence children's use of category-based induction.

Our tasks included either real items used in the training, or novel items. Our pre-test data suggested that the novel kinds had a greater perceptual lure so our finding that children made significantly more category choices for real items was as expected (Sloutsky & Fisher, 2004). However, it was interesting that there were no significant interactions between training and the type of stimuli, suggesting that the training influenced both real and novel kinds. Overall, we hypothesize that the hands-on experience with animals, along with an explanation of the importance of their non-obvious category features led to a deeper awareness of the importance of non-obvious category information.

There are three main limitations to this study: the age range, the quasi-experimental nature, and the range of stimuli used. Firstly, the age range was appropriate for the existing educational activities held at Twycross Zoo on Animal Adaptations. However, this means that our youngest group were 5- to 6-years-old, and these children already made around 50% category-based induction decisions pre-training, which limited the potential effects of training. It would be beneficial to test younger, more perceptually biased children who may benefit more from any positive effects of the training. Secondly, the study was quasi-experimental and the children were assigned to their experimental condition depending on whether their school had booked them in for an educational activity at Twycross Zoo. Unfortunately, this resulted in a non-equivalent performance of the two groups at the start of the study which made interpretation trickier: the comparison condition started with a higher category-based induction score than the training condition. Although we matched our comparison schools as closely as possible, random allocation to groups would increase confidence in our findings. Thirdly, we were restricted in the types of stimuli we could use, due to the need to manipulate category membership and similarity orthogonally. This manipulation was much easier for subordinate level category distinctions (e.g., tree frogs vs. common frogs) than it would have been for basic level distinctions (e.g., frog vs. fish). Nevertheless, it would be possible in principle to develop these types of stimuli at the basic level which would also address Gelman & Waxman's (2007) broader concerns that children are more likely to extend internal properties within the subordinate level boundaries and thus seemingly show perceptual induction when technically using subordinate-level categorical induction.

Our findings motivate further experimental work to identify the processes underlying the improvements in category-based induction we observed, specifically, the extent to which this improvement can be attributed to increased experience with real animals. Further studies could also consider factors such as learning trajectories, educational environment and taught material within the classroom.

Conclusion

We report two findings. First, in line with previous work we found an age-related increase in category-based induction inferences in 5- to 7-year-old children. We additionally found that children made more category-based induction decisions for real than novel kinds, but the effect of age group was the same across both these types of stimuli, suggesting that category-based induction develops for both over this age range. Second, we found that a one-off 30-minute interactive educational activity in which children were taught to focus on non-obvious category features increased the number of category-based induction decisions made. We would therefore conclude that even a short experience with animals, coupled with explicit teaching about the importance of category features can influence children's inductive reasoning as measured in an entirely separate testing session using computer-based induction tasks. We hypothesize that one plausible route to adult-like category-based induction is via accumulation of category knowledge through experience, which ultimately increases children's awareness of the importance of category features across the biological domain. This advances our current knowledge about the influence of biological experience on biological reasoning and highlights the potential benefits of educational activities to enhance conceptual development.

References

- Atran, S., Medin, D. L., Lynch, E., Vapnarsky, V., Ucan Ek', E., & Sousa, P. (2001). Folkbiology doesn't come from folkpsychology: Evidence from Yukatek Maya in cross-cultural perspective. *Journal of Cognition and Culture, 1*, 3–43.
- Badger, J. R., & Shapiro, L. R. (2012). Evidence of a transition from perceptual to category induction in three to nine year old children. *Journal of Experimental Child Psychology, 113*, 131-146.
- Badger, J. R., & Shapiro, L. R. (2015). Category structure affects the developmental trajectory of children's inductive inferences for both natural kinds and artifacts. *Thinking and Reasoning, 21*, 206-229.
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612-637.
- Brown, A. L. (1989). Analogical learning and transfer: What develops? In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 369-412). New York: Cambridge University Press.
- Bulloch, M. J., & Opfer, J. E. (2009). What makes relational reasoning smart? Revisiting the perceptual-to-relational shift in the development of generalization. *Developmental Science, 12*, 114-122.
- Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1147-1156.
- Coley, J.D. (2012). Where the wild things are: Informal experience and ecological reasoning. *Child Development, 83*(3), 992-1006.
- DeLoache, J. S., & Chiong, C. (2009). Babies and baby media. *American Behavioral Scientist, 52*, 1115-1135.
- Fisher, A. V., Matlen, B.J., & Godwin, K.E. (2015). Development of inductive generalization with familiar categories. *Psychonomic Bulletin & Review, 22*(5), 1149-1173.
- Gelman, S. A. (2003). *The essential child: origins of essentialism in everyday thought*. Oxford: Oxford University Press.
- Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23*, 183-209.
- Gelman, S. A., & Waxman, S. R. (2007). Looking beyond looks: Comments on Sloutsky, Kloos and Fisher. *Psychological Science, 18*(6), 554-555.
- Gentner, D. (1988). Metaphor as structure mapping: the relational shift. *Child Development, 59*, 47-59.
- Godwin, K.E. & Fisher, A.V. (2015). Inductive generalization with familiar categories: Developmental changes in children's reliance on perceptual similarity and kind information. *Frontiers in Psychology*.
- Hatano, G., & Inagaki, K. (1994). Young children's naive theory of biology. *Cognition, 50*, 171-188.
- Hatano, G., & Inagaki, K. (1999). A developmental perspective on informal biology. In D. L. Medin & S. Atran (Eds.), *Folkbiology*. (pp.321-354). Cambridge, MA, US: The MIT Press.

- Hatano, G., & Inagaki, K. (2002). Domain-specific constraints of conceptual development. *International Journal of Behavioral Development, 24*(3), 267-275.
- Hayes, B. K., Heit, E., & Swendsen, H. (2010). Inductive reasoning. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*, 278-292.
- Inagaki, K. (1990). The effects of raising animals on children's biological knowledge. *British Journal of Developmental Psychology, 8*(2), 119-129.
- Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books. a
- Lawson, C. A., Fisher, A. V., & Rakison, D. H. (2015). How children learn the ins and outs: A training study of toddlers' categorization of animals. *Journal of Cognition and Development, 16*, 236-251.
- Long, C., Lu, X., Zhang, L., Li, H., & Deák, G. O. (2012). Category label effects on Chinese children's inductive inferences: modulation by perceptual detail and category specificity. *Journal of Experimental Child Psychology, 111*, 230-245.
- Medin, D.L & Atran, S. (2004). The native mind: Biological categorization, reasoning and decision making in development across cultures. *Psychological Review, 111*(4), 960-983.
- Needham, D. R., & Begg, I. M. (1991). Problem-oriented training promotes spontaneous analogical transfer: Memory-oriented training promotes memory for training. *Memory and Cognition, 19*, 542-557.
- Piaget, J. (1964). Part I: Cognitive development in children: Piaget development and learning. *Journal of Research in Science Teaching, 2*, 176-186.
- Ross, N., Medin, D., Coley, J.D., & Atran, S. (2003). Cultural and experiential differences in the development of folkbiological induction. *Cognitive Development, 18*(1), 25-47.
- Sloutsky, V. M., Deng, W., Fisher, A., & Kloos, H. (2015). Conceptual influences on induction: A case for a late onset. *Cognitive Psychology, 82*, 1-31.
- Sloutsky, V.M. & Fisher, A.V. (2004). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General, 133*(2), 166-188.
- Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When looks are everything: Appearance similarity versus kind information in early induction. *Psychological Science, 18*, 179-185.
- Strouse, G. A., & Troseth, G. L. (2008). "Don't try this at home": Toddlers' imitation of new skills from people on video. *Journal of Experimental Child Psychology, 101*, 262-280.
- Tarlowski, A. (2006). If it's an animal it has axons: Experience and culture in preschool children's reasoning about animates. *Cognitive Development, 21*, 249-265.
- Waxman, S., & Medin, D. (2007). Experience and culture models matter: Placing firm limits on childhood anthropocentrism. *Human Development, 50*, 23-30.

Appendix

1. Hidden Properties

Sandbugs

Eats flies
 Comes from Europe
 Has cold blood
 Has soft bones
 Moves slowly
 Has a diminutive stomach

Rockbugs

Has thick blood
 Sleeps during the day
 Has an ample sized heart
 Likes to be warm
 Doesn't eat in winter
 Lays eggs

Forest weevils

Likes hot weather
 Eats wood from trees
 Is strong
 Comes from Canada
 Moves fast
 Lays 50 eggs at a time

Ground weevils

Can swim
 Only lives one year
 Communicates through vibrations
 Lays eggs in nests
 Eats grubs
 Sleeps during the winter

Tree frogs

Hunts through the night
 Have nine vertebrae
 Can be toxic
 Lays eggs in clusters
 Changes color to camouflage
 Comes from Australia

Common frogs

Has weak teeth
Doesn't drink
Has no diaphragm
Three-chambered heart
Has no ribs
Breathes through its skin when under water

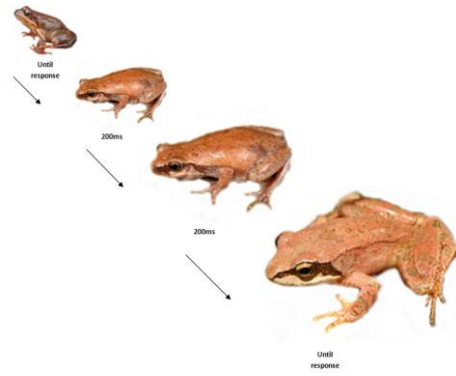
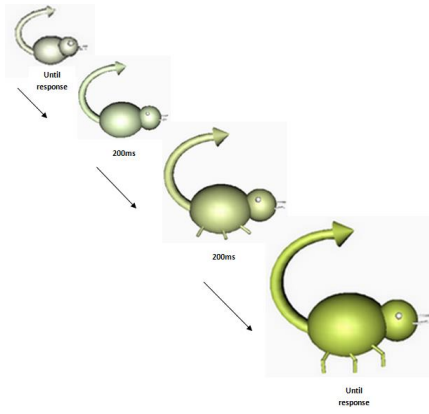
Chinchillas

Sleeps during the day
Comes from America
Lives 10-15 years
Doesn't drink much water
Has dust-baths
Doesn't smell

Squirrels

Can run fast
Is blind when young
Is active during the day
Has a great sense of smell
Can jump long distances
Sleeps during the winter

Figures



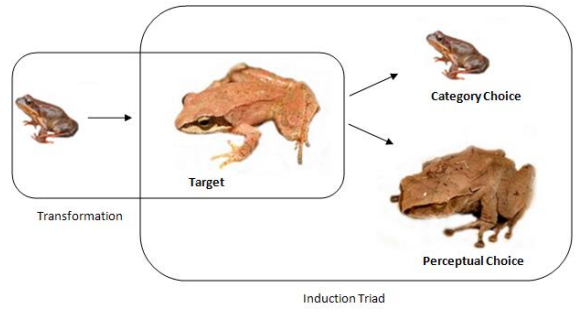
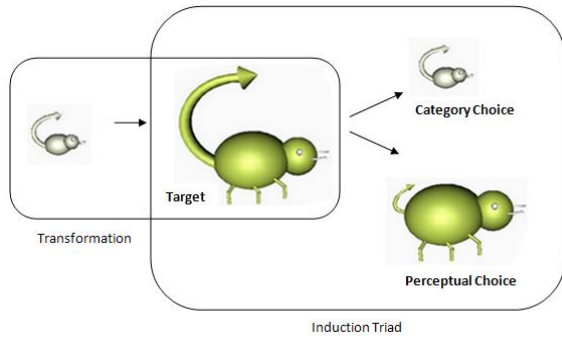




Figure 3.

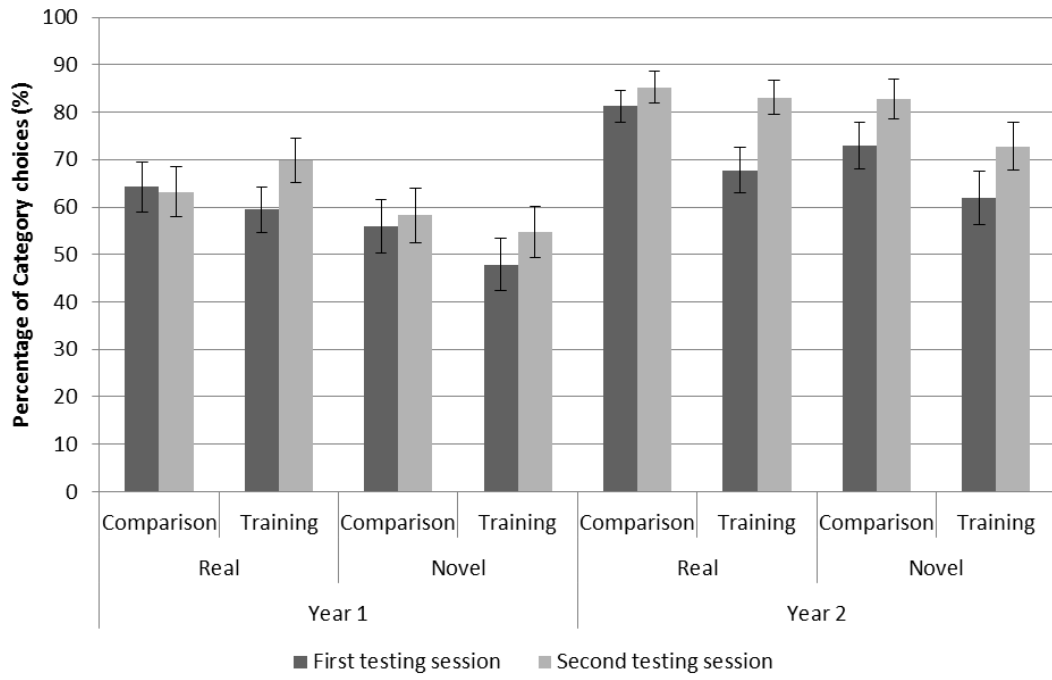


Figure 4.

Figure 1. The image on the left shows a transformation from a juvenile forest weevil into an adult forest weevil; the image on the right shows a transformation from a juvenile common frog into an adult common frog.

Figure 2. The triad on the left shows a novel kind induction triad and the triad on the right showed a real kind induction triad. The juvenile is transformed into an adult target (transformation) then the adult target is shown alongside a category choice and a perceptual choice (induction triad).

Figure 3. The animals seen during the Twycross Zoo interactive educational activities. From top left to bottom right: a White's Tree Frog, a Leopard gecko, a Madagascan hissing cockroach, a Chinchilla, and a Mexican red knee tarantula. (Images presented with permission from photographer Deborah Bardowicks, Twycross Zoo).

Figure 4. The percentage of category inferences made by children in the training and comparison conditions, split across the first and second testing sessions, for the real and novel kind tasks by year group.

Tables

Table 1. One-sample *t* tests confirming that the perceptual choice was chosen significantly more than chance for all triads directly pitting categorical choice against perceptual choice.

Type of triad (item)	Perceptual choice was chosen (%)	Statistics
Real (Squirrel)	80%	$t(30) = -161.49; p < .001$
Real (Chinchilla)	78%	$t(30) = -148.03; p < .001$
Real (Common frog)	77%	$t(30) = -149.75; p < .001$
Real (Tree frog)	77%	$t(30) = -148.03; p < .001$
Novel (Forest weevil)	95%	$t(30) = -218.00; p < .001$
Novel (Ground weevil)	95%	$t(30) = -206.67; p < .001$
Novel (Rockbug)	97%	$t(30) = -409.37; p < .001$
Novel (Sandbug)	95%	$t(30) = -225.15; p < .001$