The Two-Body Problem: Classification and Reasoning about Polymorphs

Gregory L. Murphy and Karl S. Rosengren

Four studies examined adults’ and children’s understanding of polymorphs — organisms that undergo a radical transformation (e.g., caterpillar-butterfly) — which have been used as critical tests of essentialism in concepts. In Study 1, adults judged the similarity of juvenile organisms to their adult versions and perceptual controls; in Study 2, children made forced-choice similarity judgments. Both age groups were influenced by both biological and perceptual relatedness. Age differences were largely explainable by children’s knowledge of the particular metamorphosis rather than changes in overall strategy. In Study 3, adults made inductions of different kinds of properties from juveniles to other animals; biological relatedness was most important for “deep” properties and perceptual resemblance for “superficial” properties. Study 4 tested children’s inductions on biological properties and found results paralleling their similarity judgments. These studies revealed that children and adults use both the biological relation across life stages and perceptual/morphological similarity in their thinking about these organisms. Development was largely in the direction of being better able to coordinate these two sources of information. The findings speak against accounts that emphasize perception or underlying knowledge alone.

Keywords: concepts, conceptual development, essentialism, similarity, induction

Animals that metamorphose into different life forms, polymorphs, have had an important place in the history of biological thinking. As Mayr (1982) discussed, organisms such as frogs/tadpoles, caterpillars/moths, or larvae/flies pose a problem to the traditional structural basis of biological taxonomy. Especially at higher levels of a taxonomy, the traditional (pre-evolutionary) assumption was that classes group organisms that are similar — having the same or analogous parts and structures. Polymorphs create a grave difficulty for this approach. Clearly, the tadpole is more similar to fish than it is to a salamander; but a frog is more similar to the salamander than it is to a fish. A caterpillar is more similar to a centipede or even a worm than it is to a dragonfly, but a butterfly is more similar to the dragonfly. Therefore, when biologists attempt to construct higher-level categories such as arthropods, reptiles, mammals, or anurans, the polymorphs cannot easily be categorized in a consistent way. Such problems (and others) led taxonomists to be accused of making arbitrary categories that were simply the product of their own biases about which aspects of a category were most important (Mayr, 1982).

In biology, this particular problem seems to
have been addressed by the theory of evolution, which provides a nonarbitrary way of classifying entities, namely their evolutionary relationship. However, this answer is not readily available to the layperson, who is likely unfamiliar with the evolutionary history or genetic structure of any particular organism. Laypeople may even reject the theory of evolution, which then obviously cannot serve as a basis for classification.

The psychological implications of polymorphs were perhaps first identified by Rips (1989), who developed fictional scenarios in which animals changed their forms by various means. In some cases, an animal received an environmental insult and changed its form; in others, the animal naturally developed into a new form. Rips used these scenarios to investigate people’s judgments of similarity of the different forms as well as their categorization decisions. His most notable result, which has been widely cited, was that subjects’ similarity judgments were more strongly affected by the changes the animals underwent than were categorization judgments. Rips interpreted this as revealing a difficulty with similarity-based models of categorization (see Murphy, 1993, for discussion; but see Hampton, Estes, & Simmons, 2007, for qualifications of this result).

The present investigation also looks at metamorphosis, but from a different perspective. First, we investigated familiar organisms whose metamorphosis was natural and previously known to subjects (all Rips’s examples were novel). Polymorphs are important not only as a theoretical test case but also because they are difficult for people to fit into their lay taxonomies. Understanding how people reconcile these unusual cases is necessary for a complete understanding of biological concepts. Second, rather than focusing on how a frog/tadpole is categorized, we examined how people conceive of the relation between the frog and tadpole. Although they are clearly members of the same species or even the same individual, they are also different in many respects. In particular, we examined the relative importance of perceptual resemblance and biological relatedness in people’s thinking about biological kinds. Polymorphs separate these two relationships (e.g., caterpillars do not look like butterflies but are closely related to them), and by comparing them to items that were biologically related or similar in appearance, we were able to evaluate how people use both kinds of information.

Third, we looked at processes beyond categorization, and in particular induction. It is widely agreed that induction is one of the basic functions of categorization (Murphy, 2002; Smith & Medin, 1981). The reason it is useful to identify something as a chair, for example, is because one then knows that it is suitable for one person to sit on, is made out of wood or metal, has a back, and so on, even if those properties cannot be directly observed. Simply identifying something as a category member enables one to infer likely attributes of the object. Polymorphs challenge this notion, however, because the different life forms of the organism may not give rise to the same inductions. For example, if you counted 28 legs on a caterpillar of a certain species, would you expect a butterfly of the same species to have the same number of legs? At the same time, however, the continuity of life stages suggests that there may be important commonalities across the different forms of the organism. Although body shape and behaviors may change, genetic structure remains constant, and so may some physiological properties and processes. The degree to which polymorphs enable induction is a measure of people’s weighting of underlying biological relatedness compared to other kinds of properties that differ across a polymorph’s life stages.

Finally, we examined these things in both adults and children, in order to map how the understanding of polymorphy develops. Therefore, we review the relevant adult and developmental literatures prior to presenting the experiments.
1. Theoretical Implications of Metamorphosis

One of the most important notions in the psychology of concepts is Medin and Ortony's (1989) proposal of psychological essentialism, which was in part developed to explain Rips's (1989) experiments. Medin and Ortony argued that people believe that there is an invisible, underlying essence or causal mechanism that is common to category members and that is responsible for the superficial properties of the members. Claims of actual essentialism have been largely rejected in the physical and biological sciences (Mayr, 1982). Psychological essentialism is a claim about people’s beliefs about kinds of things — not about the nature of the things themselves. For biological kinds, people usually cite genetic factors as constituting the essence; for physical elements, atomic structure; and for artifacts, function may serve as an essence (though this is controversial; see Malt & Sloman, 2007).

Psychological essentialism has been proposed to account for a number of phenomena (Gelman, 2003), but the most important ones for present purposes are demonstrations in which category membership deviates from similarity and other judgments. One such effect was reported by Keil (1989). He used transformation and discovery paradigms in which animals and objects began as one kind of thing and then were changed to or were discovered to be more like other kinds of things. For example, children were shown pictures of a raccoon that was dyed, operated on, and generally mistreated until it had the superficial properties of a skunk. Keil found that both children, after the age of 4, and adults believed that the animal was still a raccoon. Even though it looked (and smelled) exactly like a skunk, subjects agreed that its category could not be changed by such operations.

In an earlier study that involved transformations of a placid live cat via a series of highly realistic dog and rabbit masks, DeVries (1969) found that 3-year-olds generally accepted cross-species transformations. By age 4, however, children generally rejected the possibility of change across species. Similarly, Keil's youngest subjects (4 and younger) did sometimes claim that the animal's identity had changed, but even these children resisted operations when they crossed ontological categories (e.g., a transformation from animal to plant or inanimate object).

These results suggest that by age 4, even children believe that category membership is an enduring property of an object, not easily altered (see also Wells, 1930). Significantly, artifacts did not yield the same pattern of results in Keil’s (1989) study. For example, when a coffee pot was altered to look like and serve the function of a bird feeder, all age groups generally agreed that it was now a birdfeeder. This shows that children were not following a simple strategy of maintaining the initial identity, but it also suggests that because natural kinds arise from nature, they may be thought to have an essence that artifacts do not (Putnam, 1973).

2. Previous Psychological Studies of Metamorphosis

Rips's (1989) essence condition considered a natural transformation, albeit an unfamiliar one. His scenario described an animal that began life looking and acting like a bird. However, this was described as only an early stage of the animal's life; it entered a second stage in which it shed its feathers and took on characteristics of insects. Adult subjects who only read about the first stage of the animal's life rated the animal as more likely to be a bird, being more similar to a bird than an insect, and being more typical of a bird than an insect. However, subjects who read the entire scenario were more likely to categorize the animal as an insect. Their similarity and typicality ratings were also influenced by the scenario, although not nearly as much as the categorization ratings. This constituted a dissociation between categorization and resem-
blance judgments, suggesting that the biological relation of different life stages overrode the influence of perceptual similarity in categorization judgments. Rips also described a case in which a toxic mishap resulted in an animal changing from something like a bird to something like an insect. In this case, which has received much attention in the literature, people believed that the resulting animal was more similar to an insect but was in fact still a bird.

Hampton et al. (2007) followed up Rips’s (1989) results and improved the methodology in a number of ways. Although their entire pattern of results was somewhat complex, one important result was that many subjects were phenomenalists. That is, they classified the animal based on its appearance in each stage. For example, before the toxic spill, the animal was judged most likely to be a bird; afterwards, it was an insect. This result clearly is a challenge for essentialism, especially for the unnatural change, where exposure to a chemical would seem unlikely to change an animal’s essence. However, it is less clear how to interpret the phenomenalist pattern in natural metamorphoses.

Natural metamorphoses are different from the artificial kind (surgery, chemical accident), because it is in the nature of the animal to change. Real-life biological transformations exhibit characteristic organism-specific patterns controlled by underlying mechanisms (Aristotle, cited in Wiggins, 1980; Keil, 1989; Rosengren, Gelman, Kalish, & McCormick, 1991; Schwartz, 1978). In these cases, the change is part of, or caused by, the essence, rather than acting independently of it, as when there is external intervention. One goal of the present investigation is to investigate how adults and children understand the relation between the different life stages.

Developmental studies suggest that children understand some natural transformations at a relatively early age. For example, by age 3 the majority of children understand that animals get larger over the life span through growth and generally accept that animals do not shrink in size (Rosengren et al., 1991). Children’s understanding of more dramatic transformations, such as metamorphosis, is not so well established at this young age. Rosengren et al. (1991) found that most 3- and 5-year-olds expect that animals will grow in size but that their coloration or shape will remain relatively constant over the life span. Adults, in contrast to the children, were more likely to accept metamorphosis as a possible change. One interpretation of this result is that young children may assume that an animal’s essence only allows certain types of change and that these changes are more restricted than they are for adults.

In two studies designed to examine children’s reasoning about different life forms, 3-, 5-, and 7-year-olds and adults performed tasks that examined whether they expected parents and offspring to physically resemble one another (Rosengren, Taylor, & DeHart, 1997). Participants were shown a picture of an adult animal and chose the most likely offspring of the adult from four alternatives (Study 1) or picked the mother of a juvenile (Study 2). The stimuli included animals that undergo metamorphosis (butterflies, frogs), ones that vary in coloration (birds), and animals that vary mostly in size (mammals). In both studies, prior to age 7, children consistently chose on the basis of perceptual similarity, even for species that undergo metamorphosis. Adults and 7-year-olds made choices based on knowledge of life-cycle changes typical for each species. These studies suggest that young children may have a difficult time understanding the relation between different stages of life forms that undergo dramatic change such as metamorphosis, generally relying on perceptual resemblance. It is not clear from this research, however, how participants who are familiar with metamorphosis conceive of the two life stages. When the younger children did not choose the adult version as the parent, was that because they did not understand how morphology and appearance could
change across life stages, or was it simply that they did not know the specific changes that larvae or tadpoles undergo? In the present Study 2 we examine more closely how knowledge influences children’s reasoning about metamorphosis.

Finally, Keil (1992) reports a study in which children viewed pictures of novel animals and plants going through progressive stages of development. In one condition, the end state of two organisms was the same, but the earlier states differed greatly. Children were asked whether these organisms as depicted at the end were the same kind of thing. Surprisingly, children in kindergarten largely ignored the animal’s developmental progression. It was only in 4th grade (about 9 years old) that children said that the two animals were not the same more than half the time. This might seem to conflict with the studies just reviewed, as it could suggest that younger children have no particular expectations about life stages. However, Keil’s study used a conflict design, in which the identical end state of the animals suggested a “same” response, but the earlier stages suggested a “different” response. Thus, it was not designed to investigate children’s expectations of the development pattern alone. Furthermore, the use of novel animals and developmental stages may have made the task difficult for young children.

**Classification Measures**

Rips (1989) and Hampton et al. (2007) focused on similarity judgments and classification of the animals at different life stages. This is because of the important finding that people perceived some of the artificially-caused polymorphs to be more similar to one animal (the final state) but to be more likely to be another animal (the one it started out as), thereby undermining the critical assumption that similarity determines classification. Hampton et al.’s finding that this result only occurred for a minority of the subjects was therefore theoretically significant. However, if one starts out not with the goal to investigate similarity vs. classification but with the goal to understand metamorphosis — and in particular natural metamorphosis — then classification may not be the ideal measure, because of linguistic issues.

Consider a caterpillar/butterfly. As our label indicates, although we understand caterpillars and butterflies to be the same organism, we do not have a common name that refers to the animal in both of its phases. (Names for biological taxa, such as genus or species names do pick out the organism in its entire life span, but they are generally unknown to the layperson.) That is, caterpillar refers to the animal in one life stage but does not correctly refer to it in the other. One cannot point at a butterfly and say, “Look at the caterpillar” (or vice versa). This makes interpretation of classification questions somewhat difficult. If one tells a story like Rips’s, only about a familiar, naturally occurring polymorph, the questions do not quite make sense. For example, Hampton et al. asked questions such as “Before/After it changed, was the animal more likely to be a hummingbird or a bee?” In the context of an unnatural change, this question is interesting, because it might indicate whether the subject believes the animal to have changed its underlying essence, since these are normally incompatible categories. But that is not the case with normal polymorphs. Imagine that we described a Monarch and then asked, “Before it changed, was it more likely to be a caterpillar or a butterfly?” Since caterpillar refers to the life stage, the answer must be caterpillar. But this does not reflect anything about whether the animal has changed its nature or category but only the meaning of the word caterpillar, which refers to the animal in its first stage. It is much like asking, “Before/After he got out of the train, was Toshiro a passenger?” Clearly, he was a passenger before and not after, but this simply reflects the meaning of passenger, which refers to a person in a particular state, rather than classifying a person in toto. Toshiro did not change his
essence by leaving the train, even if he changed from being a passenger to a pedestrian.

Because of these problems interpreting the names of life stages, we would argue that the previous literature has not yet provided a complete understanding of natural metamorphosis. Therefore, we used somewhat different measures to investigate polymorphs, namely similarity judgments and induction. Asking whether a butterfly has the same property as a caterpillar does not ask whether a caterpillar is a butterfly, with the ensuing linguistic confusion. Furthermore, as induction is an important measure of category structure and relations, it is of interest in its own right.

3. The Current Investigation

Polymorphs and transformed entities are useful test cases for understanding the basis of concepts. However, we are concerned that past work on this topic has framed the question as a test of two incompatible hypotheses: Either people are using an essence (or deep principles of some kind) to determine category membership, or else they are using perceptual similarity. This way of framing the question is part of a long-standing concern in developmental psychology over whether children are perceptually driven, only gradually learning to use less obvious properties and principles in classification and reasoning. In this context, it is natural to construct test items that pit perceptual properties against the deeper principles, and it is an important result that the latter can overrule the former (e.g., Gelman & Wellman, 1991; Keil, 1989; Markman & Hutchinson, 1984). This way of framing the question, however, makes it difficult to discover whether both sources of information might be used, to what degree, and how these sources are coordinated (see Nguyen & Murphy, 2003).

For example, it is clear in Keil’s (1989) and Rips’s (1989) results that both perceptual and essential properties are influencing performance. For example, the mean category membership rating in Rips’s essence condition (p. 42) is about half-way between bird and insect. Theoretically, it has been argued that both perceptual similarity and underlying biological or physical relations are useful and important in identifying and reasoning about kinds (see Murphy, 2002, pp. 379–383), and Medin and Ortony (1981) argue that the essence is thought to produce the superficial properties, rather than the two being independent. Thus, the perception vs. essence contrast is not a comparison of true alternatives.

Our strategy was to test perceptual and essential properties independently. We predicted that superficial properties would in fact be more relevant for some conceptual judgments, contrary to a simplistic essentialist view. But essential properties should be more relevant for different judgments, contrary to a simplistic perceptual similarity view. To this end, we compared similarity judgments (Studies 1 and 2) with category-based inductions (Studies 3 and 4) in both children and adults. We hoped to learn from these different measures just what develops in children’s understanding of polymorphs, and why their performance in Rosengren et al.’s (1997) study deviates so markedly from that of adults. In particular, are children’s difficulties with polymorphs due to lack of specific knowledge or to a more general reliance on global similarity?

4. Study 1

In this study, adults made similarity and difference judgments of sets of animals and plants. The judgments were either within a species across different life stages or between species at similar life stages. The stimuli are listed in Table 1. The metamorphosis items included those that undergo true metamorphosis between very different life stages, as well as one plant item, a bean plant, that looks very different in its earliest life (seedling) and mature form. We included the bean in part to include plants as well as animals, and in part because our subjects were not
likely familiar with many different kinds of metamorphoses. No generalization would be gained, for example, by including many different kinds of flies and their larvae, which most subjects would not be able to distinguish. The normal items do not undergo metamorphosis, though they were chosen so that their babies were noticeably different from the adult version. For example, the infant rabbits were hairless, blind neonates, in contrast with the furry adult rabbit.

The items were triplets of labeled pictures. Each triplet included a juvenile and adult version of the organism, along with a perceptual control that resembled the juvenile. Subjects were shown a pair of items side by side (the juvenile and control or the juvenile and adult) and rated their similarity or difference. The use of two different ratings derived from a finding of Medin, Goldstone, and Gentner (1990), who discovered interesting differences between similarity and dissimilarity judgments. However, we found no differences between these measures and so don’t discuss this further.

The similarity relation has been criticized as overly context-sensitive and in some cases circular (e.g., Goodman, 1965). At the least, it is very unspecific as to what kind of properties are shared between the items judged to be similar. However, rather than being a drawback, this nonspecificity is exploited in the present experiment. We use similarity as a measure of how important people believe a given relation between the test items is. If people believe that the biological relation is important, they can rate juveniles as similar to adults; if they emphasize perceptual resemblance, they will rate juveniles as similar to their controls.

If the results are similar to those of Rips’s and Keil’s studies, we might expect to find that similarity judgments depend primarily on perceptual resemblance. As a result, the juvenile should appear to be more similar to the perceptual control than to the adult version. On the other hand, if knowledge of metamorphosis influences similarity, the adult version may be considered as most similar to the juvenile version — after all, it is the same species or, conceivably, the exact same organism. If subjects are phenomenalists, then perceptual resemblance should determine their judgments.

**Method**

**Participants.** The participants were 32 University of Illinois undergraduates who participated for pay or for course credit.

**Procedure.** Eight triplets were constructed (shown in Table 1) and illustrated with photographs or highly realistic colored illustrations.

---

**Table 1** Items used in the similarity and induction tasks.

<table>
<thead>
<tr>
<th>Metamorphosis Items</th>
<th>Perceptual Control</th>
<th>Adult</th>
</tr>
</thead>
<tbody>
<tr>
<td>Juvenile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>monarch caterpillar</td>
<td>gypsy moth caterpillar</td>
<td>monarch butterfly</td>
</tr>
<tr>
<td>bean seedling</td>
<td>beech seedling</td>
<td>bean plant</td>
</tr>
<tr>
<td>tadpole</td>
<td>cod</td>
<td>frog</td>
</tr>
<tr>
<td>housefly larva</td>
<td>dragonfly larva</td>
<td>housefly</td>
</tr>
<tr>
<td>Normal Items</td>
<td></td>
<td></td>
</tr>
<tr>
<td>baby owl</td>
<td>baby seagull</td>
<td>owl</td>
</tr>
<tr>
<td>border collie pup</td>
<td>terrier pup</td>
<td>border collie</td>
</tr>
<tr>
<td>infant bunny</td>
<td>mouse pup</td>
<td>rabbit</td>
</tr>
<tr>
<td>baby alligator</td>
<td>chameleon</td>
<td>alligator</td>
</tr>
</tbody>
</table>
Table 2 Results of Similarity and Different Ratings, Study 1. Both ratings were on a 1-7 scale.

<table>
<thead>
<tr>
<th></th>
<th>Normal items</th>
<th>Polymorphs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Juvenile-Adult</td>
<td>5.3</td>
<td>3.9</td>
</tr>
<tr>
<td>Juvenile-Perceptual Control</td>
<td>4.4</td>
<td>4.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Normal items</th>
<th>Polymorphs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Juvenile-Adult</td>
<td>2.5</td>
<td>3.7</td>
</tr>
<tr>
<td>Juvenile-Perceptual Control</td>
<td>3.9</td>
<td>3.8</td>
</tr>
</tbody>
</table>

Note. High similarity ratings correspond to low difference ratings.

Each triplet yielded two test items, the juvenile-adult pair and the juvenile-control pair. Each picture was labeled so that there would be no ambiguity about the item’s actual identity (subjects may not be as familiar as they ought to be with specific kinds of larvae, for example). The pictures were divided up into two booklets such that if the juvenile-control pair appeared in one booklet, the juvenile-adult pair appeared in the other booklet. Subjects rated the items from both booklets, with the order counterbalanced. Half the subjects rated the similarity of the pairs of pictures and half rated their dissimilarity, both on a 1-7 scale.

5. Results and Discussion

Table 2 shows the mean ratings of similarity and dissimilarity. Note that high similarity ratings correspond to low difference ratings. The results are easy to summarize. In every case, the ratings were fairly similar, falling in the middle of the range, except for the normal juvenile-adult normal, which were rated as more similar and less dissimilar. One way to explain these results is that there are two components to subjects’ ratings: perceptual resemblance and underlying biological relatedness. The perceptual controls were perceived as somewhat similar to the juvenile, because they look the same, and the adult metamorphosis items were somewhat similar because they are the same organism. The normal adult items were perceived as even more similar because they have both forms of resemblance — they look like the juvenile, and they are the adult version.

A two-way within-subjects ANOVA was performed to analyze each dependent measure. In the similarity ratings, the polymorphs were rated as less similar than the normal items, $F(1, 15) = 17.40, p < .001$, but this was mitigated by the interaction, $F(1, 15) = 10.18, p < .01$, reflecting the fact that the normal adult was rated the most similar (and the polymorph adult least). In the difference ratings, the pattern is even more clearly one of a single condition (adult normal) being different from the others. As a result, both main effects were reliable, $F(1, 15) = 9.03, p < .01$ for the advantage of normal items, and $F(1, 15) = 5.5, p < .05$ for the adult vs. perceptual control comparison, as well as the interaction, $F(1, 15) = 13.55, p < .01$.

These findings do not support the idea that similarity is more sensitive to perceptual resemblance than to underlying biological relations. In the metamorphosis items, the (biologically related) adults were judged to be equally similar to the juveniles as the (perceptually related) controls were. Thus, although perception is important, similarity ratings cannot be characterized as purely “phenomenalist,” as biological relatedness is also a factor.

Because people are familiar with only a few
forms of metamorphosis, the number of items used was relatively small. We had no particular predictions about how the specific items might differ, but we examined them to see if there was any pattern evident in the results. All four of the normal items showed the expected pattern of greater similarity to the adult than to the perceptual controls. For the metamorphosis items, this pattern was reversed for the caterpillar and fly larva. It seems likely, and not surprising, that this is due to the greater changes from juvenile to adult forms for these items. In both cases, the animal changes from a rather undifferentiated worm-like creature to a flying animal with distinct body parts. Thus, the amount of change in the metamorphosis seems to influence similarity/difference judgments. Another possibility is that familiarity with the metamorphosis influenced results (e.g., the tadpole-frog metamorphosis is very familiar, and these items were rated as much more similar than their perceptual controls). We explore this possibility in the next study, where we obtained specific information on children’s knowledge of the metamorphosis.

In summary, the results show that the metamorphosis relation is known and used by college students in their judgments of similarity: although the life forms look different, they were rated as similar as items that resemble each other (but are not related). However, the results do not suggest that the identity of an organism is the sole factor in similarity judgments. That is, even though a tadpole is a frog, frogs are no more similar to tadpoles than codfish are. Normal adults were perceived as more similar to their juvenile forms than polymorph adults were, suggesting that both perceptual and biological relations were important in judging similarity of polymorphs.

6. Categorization Posttest

Although not the focus of our research, it was of interest to know whether Rips’s (1989) result showing a difference between categorization and similarity judgments would be found with familiar organisms that actually undergo metamorphosis. The perceptual control and the adult versions were both moderately similar to the target juvenile organism, yet categorization seems to clearly favor the adult version. That is, the dog and puppy are more similar than the frog and tadpole are, yet both pairs should be classified together. Thus, similarity ratings do not appear to predict categorization judgments.

This reasoning relies on our intuitive argument that the different life forms would be classified together. Although it seemed clear to us that subjects would categorize the tadpole as a frog, rather than a cod, we verified our intuitions in a short task with 18 subjects. (One other subject was discarded for giving perverse responses to the normal items, perhaps misunderstanding the task.) The subjects were shown the labeled juvenile picture and made a forced choice categorization judgment between the perceptual control and adult categories, by circling one of the names on a response sheet. That is, they were asked to say whether the tadpole should be categorized as a cod or a frog.

For the normal items, the adult form was chosen 97% of the time. For the metamorphosis items, the adult form was chosen 83% of the time, which was significantly less often, $t(17) = 2.56, p < .05$. Although subjects said that the juvenile polymorphs were in the same category as the adults the large majority of the time, they equally clearly did not do so at ceiling. The caterpillar and fly larva showed the lowest level of categorization (67% and 78%, respectively), and the frog and bean plant showed higher levels (100% and 89%). This parallels the similarity results in Study 1, where the frog and bean plant had the greatest similarity between adult and juvenile forms of the metamorphosis items.

In short, our results did not document a dissociation between similarity judgments and category membership, because the metamorphosis organisms were both less similar to their juve-
nile forms and less likely to be categorized in the same way. Thus, these results provide a contrast to Rips’s (1991) experiments, which found a dissociation, and are consistent with the general trend of Hampton et al.’s (2007) results, which did not. However, our design was somewhat different, as we compared categorization of the juvenile to its adult form vs. a different, perceptually similar animal. Rips did not use such a perceptual control. Finally, we should emphasize that, in general, people overlooked the differences between the metamorphosis life stages and classified them together. That is, the majority were not phenomenalists.

7. Study 2

Study 2 used a modified procedure to investigate children’s similarity judgments of the same metamorphosis pairs rated by college students in Study 1. Rather than asking children to rate pairs on a 7-point scale, which they would likely find difficult to do, children were asked to pick which of two items either were “most like” or “had the same insides.” These questions enabled us to look more closely at whether children rely on perceptual similarity, morphological differences, or biological relatedness when reasoning about different life forms of the same species. Gelman and Wellman (1991) used similar questions to examine children’s reasoning about nonobvious properties. They presented children with triads consisting of a target item and two test items. One test item was from the same category as the target but varied in perceptual appearance. The other test item was from a different category as the target but was similar in appearance. Children were asked “Which of these looks most like the [target]?” and “Which of these has the same insides as the [target]?” Both 3- and 4-year-olds were above chance in answering based on either perceptual resemblance or biological relatedness. We used the form “is most like” rather than “looks most like” as Gelman and Wellman did, so that children could decide for themselves what aspect of the items (perceptual, biological, or whatever) to use in deciding similarity.

The current investigation examined how children judge the similarity of polymorphs. If the results are similar to those of Rips’s and Keil’s studies, we might expect children to respond to the “most like” question primarily based on perceptual resemblance. If so, the juvenile should be judged to be more similar to the perceptual control than to the adult version. On the other hand, if knowledge of metamorphosis influences similarity, the adult version may be considered as most similar to the juvenile version — after all, it is the same species. We included the “same insides” question because this question can be interpreted somewhat differently for polymorphs than for the normal items Gelman and Wellman tested. For organisms such as caterpillars and butterflies, or tadpoles and frogs, internal structures can vary across the different life forms, so they may be viewed as not having the same insides. However, this question can also be viewed as tapping into deeper biological relatedness. Choice of the perceptual control would indicate reasoning based on similarity of internal physical structures, while choice of the adult form would indicate reasoning based on deeper biological relatedness. At the end of the session, we assessed children’s knowledge of the life stages of the polymorphs.

Method

Participants. Forty children (23 male) from local preschools ($M$ age = 4;6; range 4;2 to 5;8), 43 children (20 male) from local public first grade classrooms ($M$ = 6;11; range 6;0 to 7;6), 27 children (15 male) from local third grade classrooms ($M$ = 9;2; range 8;0 to 9;8), 18 children (10 male) from local fifth grade classrooms ($M$ = 10;4; range 10;0 to 11;10) and 27 University of Illinois undergraduates ($M$ = 21) participated.

Procedure. The same eight triplets from Study 1 were used in this study (see Table 1). As in
Table 3  Percentage choice of same-species items for most-like and same-insides judgments, Study 2.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Most-Like Judgments</th>
<th>Same-Insides Judgments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal items</td>
<td>Polymorphs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 to 5</td>
<td>78</td>
<td>23</td>
</tr>
<tr>
<td>6 to 7</td>
<td>40</td>
<td>55</td>
</tr>
<tr>
<td>8 to 9</td>
<td>60</td>
<td>58</td>
</tr>
<tr>
<td>10 to 11</td>
<td>40</td>
<td>53</td>
</tr>
<tr>
<td>Adults</td>
<td>95</td>
<td>88</td>
</tr>
</tbody>
</table>

the previous study, the pictures were presented in a loose-leaf booklet. However, here the left page of the booklet contained a labeled picture of a target item (e.g., monarch caterpillar) and the facing page contained a perceptually similar juvenile item (e.g., a gypsy moth caterpillar) and the adult version of the target item (e.g., a monarch butterfly). Children were presented with the eight triplets in one of two random orders. When the items were presented, each was labeled and the children were asked to point to each item (e.g., the experimenter asked: “Which one is the monarch butterfly? Which is the gypsy moth caterpillar? Which is the monarch caterpillar?”). For half of the children (most like condition), the experimenter pointed to the target item on the left-hand page and then asked “Which of these two [pointing to the items in turn] is most like this one [pointing to the target]?” For the other half of the children (same insides condition), the experimenter asked “Which of these two has the same insides as this one?”

After all eight triplets had been completed, we assessed each child’s knowledge of life cycle changes. Children were first shown pictures of a juvenile and adult horse and were told “Here is a picture of a horse and here is a picture of a baby horse. When this baby horse grows up it will be just like this horse.” Children were then shown pictures of adult versions of an owl (as a warm-up) and the frog, house fly, butterfly, and bean plant. The children were asked to describe what each one was like as a baby or where it came from. Each response was coded by the two authors (who were blind to the children’s similarity judgments) as demonstrating (1) or lacking (0) clear knowledge of juvenile life forms. The resulting knowledge score ranged from 0-4, indicating the number of metamorphoses the child was familiar with.

Results

Table 3 presents the mean judgments in percentages for the most like and same insides questions. Scores close to 100 correspond to a choice of the different life stages of the same species and scores closer to 0 correspond to choices of the perceptually similar control item from a different species. As can be seen in the table, participants generally chose the same species item more often for the normal animals compared to the polymorphs for both question types. However, considerable variability is evident in the pattern of results especially among the intermediate age groups.

We analyzed the percentage of same-species responses in a 2 (change type: normal, polymorph) × 2 (question type: most like, same insides) × 5 (age group) ANOVA with change type as a repeated measure. Participants generally chose the same species item more often for the normal items ($M = 72$) than for the polymorphs ($M = 49$) ($F(1,144) = 47.7$, $p < .001$). However, this difference was much greater when participants were asked the same insides question ($M$ difference = 35) compared to the most like question ($M$ difference = 7) (interaction $F(1,144) = 21.6$, $p < .001$). There was no overall effect of the ques-
tion type, however \( F(2, 144) = 1.9, p > .15 \).

Generally, with increasing age there was a tendency to choose the same-species item \( F(4, 144) = 2.2, p = .075 \), but this was mitigated by a question type by age interaction \( F(1, 144) = 5.8, p < .001 \). Children were slightly more likely to choose the same species animal for the same insides question \( (M = 65.3) \) than for the most like question \( (M = 59) \), but adult subjects exhibited the opposite pattern, choosing the same species items significantly more often for the Most Like question \( (M = 91.5) \) compared to the Same Insides \( (M = 56) \). A significant age by item type interaction was also obtained \( F(4, 144) = 9.5, p < .001 \). For the polymorphs, with increasing age, participants were more likely to choose the same species item rather than the perceptual control \( (M = 32 \text{ for 4-5 year-olds to } M = 61 \text{ for adults}) \). For the normal items, the same species items were chosen quite often by both the youngest children \( (M = 83) \) and the adults \( (M = 86) \), with children in between these ages choosing the same species less often.

Finally, a significant three-way interaction was obtained between age, question type and animal type \( F(4, 144) = 4.9, p = .001 \). This interaction reflects idiosyncratic differences between the different age groups that do not seem obviously interpretable \( \text{(e.g., the 8-9-year-olds’ higher same-species choice than the younger and older age groups for the normal items in the most-like condition; the higher 10-11-year-olds’ proportions for polymorphs for the same-insides question)} \). However, this interaction does not contradict the larger effects reported above, most importantly the greater proportion of same-species choices for the same-insides question and the greater difference between polymorphs and normal items for the same-insides question.

To explore whether knowledge of the life cycle of different animals influenced the pattern of results we repeated the analysis but with knowledge score entered as a covariate. Overall the same pattern of results was obtained, except that the age trend was no longer even marginally significant \( F(4, 142) = 0.9 \). Instead, there was a significant main effect of knowledge \( F(1, 142) = 9.9, p < .01 \) reflecting that subjects with high levels of knowledge were more likely to choose the same-species items than were those with less knowledge.

**Discussion**

The large number of interactions in the ANOVA may make the results difficult to grasp; some of them are probably due to idiosyncratic differences in the classes tested. The following effects are the most theoretically important. First, when determining which item was most similar to the target, participants were more likely to choose the different life stage of the same species for normal items than for the polymorphs. This tendency was stronger when participants were presented with the same-insides question. Thus, many participants—and especially adults—think that an animal’s “insides” change with metamorphosis.

The youngest children were somewhat different from the older groups. They chose the same-species item for normal animals much more than they did for the polymorphs. That is, a puppy was viewed by the 4-5-year-olds as more like the adult dog than it was to a similar looking puppy from a different species. But for polymorphs, the youngest children generally chose the perceptual control over the adult life form. The response reversal across these two different types of items implies that the children considered biological relatedness for normal items but not for polymorphs. That is, the parental relationship was able to overcome the mild perceptual conflict for dogs, bunnies, and so on. But for polymorphs, the biological relationship could not overcome the striking differences in perceptual resemblance. The analysis using knowledge as a covariate eliminated the main effect of age, suggesting that the youngest children differed from
the others primarily in being less familiar with the metamorphoses.

Adult subjects exhibited a distinctly different pattern, agreeing with the younger participants in choosing the adult form of the same species for the normal items, but also choosing the adult form for the polymorphs in the most-like judgments. This implies that adults are making this judgment based on biological relatedness for both types of items. However, for the same-insides questions, adults surprisingly revert to choosing the same-aged item for the polymorphs (e.g., the dragonfly larva went with the housefly larva). This is unexpected under the assumption that “same insides” is a measure of underlying biological relations. However, as we will discuss, this response is quite reasonable under slightly different assumptions.

The pattern of results in the intermediate ages is less clear. Many of their responses lurk in the intermediate 40-60% range, suggesting no clear preference for either type of item. These ages also include the cells (mentioned above) that seem out of line with the other groups’ responses (e.g., higher proportions for 10-11-year-olds for polymorphs in same-inside judgments). These patterns may reflect the interaction between the children’s everyday reasoning about biology and specific facts learned in school biology (Hatano & Inagaki, 1987). For example, children in one class may know about frogs and tadpoles and choose the same-species picture fairly often for that item, but not know about larvae or plants and so choose the perceptual controls consistently for them. One implication of this result is that when children learn about one or two polymorphs, they may not generalize their reasoning for these items to other forms that they aren’t as familiar with.

The youngest children in our study predominantly chose the different life form of the same species over the perceptual control for the normal items in response to the same-insides question. Given that these children generally chose the perceptual control for this question for the polymorphs, this choice suggests that the youngest children are basing their response on shared physical structures rather than on deeper biological relatedness. That is, the fact that animal A grows up to be animal B does not by itself entail for them that the animals have the same insides, if their morphologies are very different. The same pattern was exhibited by most of the other age groups including the adults. Our oldest children, 10- to 11-year-olds, exhibited a slightly different pattern, choosing the different life form of the polymorphs for both the same-insides and most-like questions. It is tempting to interpret this pattern as a rigid application of the biological relation: If A grows into B, then they are the same, and so should have the same insides, but further investigation is needed to confirm this interpretation.

Taken together, the results suggest that with increasing age there is an increased tendency to make judgments based on deeper biological relatedness. However, if the different life forms of a species vary greatly in appearance, as they do in the case of polymorphs, participants switch from using deeper biological relatedness, such as genetics or parentage, to responding on the basis of shared physical structures.

Part of what we are suggesting is that the same-insides question used by Gelman and Wellman (1991), among others, should not be viewed as a pure measure of an underlying essence. For example, even college students judged that different life stages of polymorphs don’t have the same insides, presumably because they recognize that the internal structures needed to support, say, many pairs of legs and massive eating are different from those necessary to support flying, migrating, and mating. So, even though a monarch caterpillar and butterfly may be genetically identical, they do not have identical insides. But this question was not answered purely by virtue of perceptual similarity, because the baby owl and baby bunny look much more simi-
lar to their different-species matched babies than they do to their own grown-up version, yet children and adults chose the adult animals for the same-insides question. Thus, the phrase “same insides” may be indicating internal morphological and physiological structures that are likely common across the life span (for normal animals) rather than parentage, genetic structure, and other micro-biological properties. Thus, this question is probably not a clear measure of an underlying essence.

Finally, it is worth noting that children and adults showed considerable flexibility in their use of categorical and perceptual information. Most age groups showed a significant difference between the two questions (compare left and right sides of Table 3), revealing that they were able to switch responding based on the particular question being asked. This seems least true for the 4-5-year-olds, who gave fairly similar responses to the two questions. The largest change was for college students, who picked the same species for the polymorph item for the most-like question 88% of the time, dropping to 35% of the time for the same-insides judgment. So, at least by age 6, children can choose different items as most similar to target, depending on the specific question asked, rather than relying on a fixed similarity relation (see Nguyen & Murphy, 2003).

8. Study 3
Perception of similarity is not generally sufficient to determine category membership, as numerous studies have found (e.g., Keil, 1989). If similarity were the only determinant of categorization, then people would often classify a cod and a tadpole together, since they are just as similar as the tadpole and frog. However, people do not think that cods and tadpoles are the same kind of thing; they know that tadpoles become frogs. An important question is to what degree that relationship is informative about an organism’s properties. For example, if people learn that one robin uses a certain digestive enzyme, they are likely to generalize that fact to other robins and are fairly likely to generalize it to birds as a whole (Gelman & Markman, 1987; Osherson, Smith, Wilkie, López, & Shafir, 1990; Rips, 1975). It is this ability to make inductions across category members that makes categories so useful. But does inductive generalization also hold across the life forms of polymorphs?

From one perspective, the answer to this question clearly should be “yes.” Because a tadpole and a frog could be the exact same individual, learning a property of one must be relevant to judging the property of the other. If a tadpole uses a certain digestive enzyme, one might expect that the enzyme does not disappear in the adult form. From a different perspective, however, the answer to the question is not so clear. A tadpole has a long tail and no legs; a frog has legs and little if any tail. The behaviors, ecological niches, and even body parts of different life forms can be quite different. Clearly, some important properties of the juvenile version do not hold for the adult version of the organism, and so induction is thereby limited. Responses to the “same insides” questions of Study 2 show that adult subjects do not assume that polymorphs retain hidden properties across metamorphosis. This raises the question, when the animal appears to change utterly, as in the caterpillar-butterfly example, will people draw inductions from one life stage to another?

Findings on category-based induction show that not all inductions act alike. As Gelman and Markman (1987) first noted, some properties are inducible across category exemplars (e.g., for animals, basic behaviors and internal structures), whereas others are not (e.g., dirtiness or size). A number of researchers have demonstrated that strength of induction depends greatly on the property being tested and its relation to the category members (Heit & Rubinstein, 1994; Kalish & Gelman, 1992; Nguyen & Murphy, 2003; Ross & Murphy, 1999). For example, if two animals are related biologically, bi-
ological attributes may be more inducible than behavioral attributes, and vice versa if the animals are related in terms of their habitat (Heit & Rubinstein, 1994). If one knows multiple attributes about a category member, the causal or relational structure among the attributes may influence whether one induces one of them to a new category member (Lassaline, 1996). Although formal models of category-based induction have relied almost completely on overall similarity to predict induction (Osherson et al., 1990; Sloman, 1993), this research suggests that induction may be achieved by a more complex reasoning process, in which subjects decide whether the critical feature is a plausible property of the new item, given its other properties (Johnson & Keil, 2000; Proffitt, Coley, & Medin, 2000).

This literature suggests that induction in polymorphs may well depend on the nature of the property being induced. Above we suggested that properties such as using a digestive enzyme might continue across the metamorphosis, whereas other properties would not. To examine this possibility, we tested three kinds of properties. One kind was internal biological properties, such as “has 12 chromosomes” and “has sesamoid bones.” These properties seemed to be the kinds of things most likely to be preserved across metamorphosis. Another kind of property was the superficial behavioral type, which described a behavior (loosely speaking, when applied to the plant) or predisposition that seemed to arise from the shape, size, and body parts of the animal. For example, “moves around mostly by crawling” and “dies from exposure to below-freezing temperatures” were two such cases. These properties were designed so that they were physically possible for all members of a triad but would be most likely to generalize from the juvenile items to their controls. For example, if a bluebottle fly larva moves by crawling, then so does another larva; if an infant bunny (which was naked in the picture) dies from below-freezing temperatures, then so would a naked mouse pup. It is less clear that such properties would generalize across the different forms of the polymorphs.

Finally, we asked whether the (expected) difference between such properties was due to the reasoning process we suggested or instead was a simple function of biological vs. behavioral properties. Perhaps behavioral properties simply do not generalize as well. Therefore, we developed a list of deep behavioral properties (again using the term behavioral loosely) that were more likely to be species-specific behaviors or predispositions than to derive from body configurations and superficial properties, such as “very active after dark,” and “grows best in hot climates.” (We should note that our classification of deep and superficial properties is relative to the particular category under consideration. We do not claim that a given property would always have this classification — only that for the category it was tested on, each property was based on superficial or deep causes. Eating bugs may be a deep behavioral property for a frog but not for a game-show contestant, for example.) The results will serve as a manipulation check on our categorization of the different properties.

The experiment asked three basic questions. First, would there be any substantial induction from juvenile to mature forms of the polymorphs? When an animal changes size, shape, and behaviors drastically, it is not clear that any properties will be generalized from one life form to the other. Second, would such induction be segregated according to property type? We suspected that the biological and deep behavioral properties would be generalizable for the polymorphs but that superficial behavioral properties would not. And third, would some inductions (the superficial ones) be stronger to control items than to the same animal in a different life stage? In keeping with Heit and Rubinstein’s (1994) results, we suspected that the tadpole and cod would be perceived as sharing superficial behavioral features more than would the tadpole and frog, even though the latter pair have a much
stronger biological relation than the former pair.

We investigated these questions by telling subjects that one property was true of the juvenile form of one of the items and then asking them to rate how likely the property was to be true of the test item — either the control or adult form.

**Method**

**Subjects.** The subjects were 46 University of Illinois students.

**Materials and Procedure.** The same triplets used in Experiment 1 served as stimuli here. For each triplet, three properties were constructed as described above: a biological property, a superficial behavioral property, and a deep behavioral property (listed in the Appendix). Picture booklets similar to those used in Experiment 1 were again used here. On each pair of facing pages appeared two labeled pictures: the juvenile organism and either the control or the adult form. Subjects read the induction question on a questionnaire, answered it, and then turned the page of the booklet for the next question. Each subject went through two booklets. The booklets differed in which item was paired with the juvenile. If the control appeared in booklet 1, then the adult form appeared in booklet 2.

A cover sheet introduced the notion of induction. Subjects were told that they were to assume that the property given was true of the first item and then to provide an estimate of how likely the second item was to have the property on a 0% to 100% scale. The questions followed the form: “Suppose you knew that the baby owl has sesamoid bones. How likely do you think it is that the baby seagull has sesamoid bones?” (The pictures of the two named items were in view while the question was answered.) Subjects wrote their responses directly below the question. The questionnaire appeared in four different forms. Forms 1 and 2 (completed by 24 participants) contained the biological and superficial behavioral properties, whereas forms 3 and 4 (completed by 22 participants) contained the biological and deep behavioral properties. The difference between forms 1 and 2 and between forms 3 and 4 was simply whether the biological or behavioral question was asked about each picture pair. For a given subject, then, the baby owl would appear twice — once paired with the adult owl and once with the control (the baby seagull). Furthermore, the subject would answer one biological and one behavioral question about the baby owl. As a result, an individual subject was never asked two questions about the same pair of items, nor was a question ever repeated.

**Results**

**Qualitative results.** Table 4 shows the pattern of results across property and item type. Due to the complexity of the design, we begin by making qualitative comparisons and report the supporting statistical analyses afterwards. (Also, some of the most important results are whether there was substantial induction on an absolute scale, which is not revealed in an ANOVA.)

First consider the normal items, such as the rabbit and owl, shown in the first two columns. As expected, the biological properties were strongly generalized to the adult version of the animal, more strongly than they were to the perceptual controls. (The two entries for biological properties are different counterbalancing groups answering the same questions, so these are just replications of the same result.) The deep behavioral properties were also more strongly generalized to the adult organism than they were to the perceptual control. (The two entries for biological properties are different counterbalancing groups answering the same questions, so these are just replications of the same result.)
Table 4 Mean induction ratings in Study 3.

<table>
<thead>
<tr>
<th></th>
<th>Normal Items</th>
<th>Polymorphs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Perceptual Control</td>
<td>Adult</td>
</tr>
<tr>
<td>Biological Features</td>
<td>46</td>
<td>80</td>
</tr>
<tr>
<td>Deep Behavioral Features</td>
<td>34</td>
<td>80</td>
</tr>
<tr>
<td>Biological Features</td>
<td>55</td>
<td>79</td>
</tr>
<tr>
<td>Superficial Behavioral Features</td>
<td>73</td>
<td>17</td>
</tr>
</tbody>
</table>

Note. The two sets of biological features are the same questions answered by different subjects. The first set was answered by subjects who also rated the deep behavioral features, and the second set was answered by subjects who also rated the superficial behavioral features.

types.

The main question of interest, then, is how these results compare to those of the polymorphs. As Table 4 shows, the biological properties did have fairly strong inductive strength from the juvenile to the adult form. Across the two replications, the mean likelihood was 71%, which was 16% greater than the induction to the perceptual controls. This level of inductive strength was not quite as strong as that found in the normal items, however.

The behavioral properties also performed similarly with the polymorphs as they did with the normal items. The deep behavioral properties were rated as 33% more likely for the adult than for the perceptual control. In contrast, the superficial properties were rated as 36% more likely for the perceptual control. In short, subjects were very willing to draw inductions across the life stages of polymorphs, when the properties were of the right sort. However, there is a suggestion that such inductions were weaker than those drawn across the life span of normal animals that do not undergo metamorphosis. Not surprisingly, people believe that there is more commonality between a baby and adult owl than they believe there to be between a larva and fly. However, in spite of the massive change in life form, subjects found important commonalities between the stages of the polymorphs, for properties that relate to deeper biological properties or basic behavioral predispositions. Subjects neither attributed all properties across the lifespan, nor did they decide that that no inductions were likely, given the large differences. Instead, they distinguished properties that are and are not generalizable across life form.

Formal analyses. To analyze the data more formally, we separated the two groups of subjects so that the first two rows in Table 4 were analyzed separately from the last two rows. This way, the behavioral-biological feature comparison is manipulated purely within subjects. (Combining the two groups would have resulted in an awkward analysis, because they answered the same biological questions but different behavioral questions, making the analysis not purely within- or between-subjects.)

The first analysis, for subjects who received the deep behavioral features (top two rows of Table 4), revealed that induction was much stronger from the juvenile to the adult than to the perceptual control, $F(1, 21) = 80.15$, $p < .001$. This was especially true for the behavioral properties, interaction $F(1, 21) = 11.16$, $p < .01$. There was also a marginal advantage for the biological question overall, $F(1, 21) = 4.15$, $p < .06$. There was also an interaction of organism type (polymorph vs. normal) and comparison item (perceptual control vs. adult), $F(1, 21) = 8.45$, $p < .01$. The polymorphs showed less of an advantage for the adult over the control item than the normal items did. This is consistent with the similarity results in Study 1, where polymorph
juveniles were moderately similar to both their adult versions and perceptual controls. That is, not surprisingly, induction from a baby rabbit to an adult is stronger than induction from a larva to a fly. No other factors approached significance.

The second analysis, for subjects who received the superficial behavioral features (bottom rows of Table 4), found an advantage of induction to the perceptual control over the adult, opposite to the effect with deep features, $F(1, 23) = 9.54$, $p < .01$. The difference was opposite here because the perceptual controls had strong inductions for the superficial behavioral property. The analysis again found stronger inductions for biological properties over the behavioral ones, $F(1, 23) = 27.89$, $p < .001$. There was also a highly reliable interaction of question type and comparison item, $F(1, 23) = 183.09$, $p < .001$, reflecting what is apparent in Table 4 — that superficial behavioral features were induced to control items, but biological features were induced to the adult form. Finally, there was a barely significant 3-way interaction of all the variables, $F(1, 23) = 4.62$, $p < .05$, reflecting the fact that the just-reported 2-way interaction is stronger for the normal items than for the polymorphs. This seems to reflect a slight vagueness in the polymorph inductions: All of the means are closer to 50% for the polymorphs than for the normal items.

In a number of places we have alluded to the perception that induction was somewhat weaker in the metamorphosis than in the normal items. That claim is reflected in a number of effects, but none of them is a very pure test of the effect, due to the interactions involving feature type. Therefore, we performed one final analysis that included only biological features — i.e., rows 1 and 3 in Table 4. If induction is weaker to the polymorphs, then there should be an interaction of organism type and comparison type (the perceptual controls serving as a control for any overall preference for baby rabbits over tadpoles and the like). The analysis did in fact find the expected interaction, $F(1, 45) = 4.33$, $p < .05$: The advantage of adults over perceptual controls was 29% for the normal items, but only 16% for the metamorphosis items. A separate test of the 16% advantage within the metamorphosis items showed that it was highly reliable, $F(1, 45) = 13.28$, $p < .001$. Thus, even polymorphs do show an effect of biological relatedness, but the effect is not as great as that for organisms that do not undergo metamorphosis and therefore resemble their adult form more closely.

9. Discussion

The results provide valuable information about the coherence of metamorphosis categories. First, it is important that subjects did in fact draw inferences across life forms of the polymorphs. In spite of sometimes large differences in size, morphology, behavior, and habitats, a novel fact learned about the juvenile was generalized to the adult. Second, this induction was generally greater than that to perceptual controls. Thus, subjects viewed the biological relationship of a larva and fly as being more informative than that of two larvae of different species, in spite of their greater perceptual similarity.

The results also revealed inductive flexibility, as in past studies of induction (Heit & Rubenstein, 1994; Kalish & Gelman, 1992; Nguyen & Murphy, 2003; Ross & Murphy, 1999). Although induction was strong from juvenile to adult for biological and deep behavioral properties, it was not strong for properties related to superficial aspects of the stimuli. One should not take the word superficial too literally here, as this simply means properties shared with the perceptual control stimuli and could refer to fairly important aspects of the organisms (presence of fur, for example). Nonetheless, the interactions involving property type indicate that only some kinds of properties can be induced from one stage of a polymorph to another. However, the same is also true of the normal items (see last line of Table
4). Some properties of baby owls are not true of adult owls, and vice versa. The babies can’t fly and are helpless; the adults are skilled fliers and hunters.

Heit and Rubinstein (1994) compared behavioral and underlying biological properties, finding that, for example, a whale and rabbit might share the biological property, but the whale and a tuna the behavioral property. Thus, they argued that behavior is more related to superficial property overlap, whereas biological properties are more related to underlying relations. Our results qualify their findings, as it is possible to find behavioral or dispositional properties that reflect deep biological relations and hence are induced across ages and life forms. We are not questioning their results, but we suggest that behavioral vs. biological properties may not always correspond to surface vs. deep properties.

The final important finding is the overall result that induction seems to be generally weaker for the polymorphs than for the normal organisms. Even the biological properties, which were designed to be shared across life forms, were not induced as strongly for the polymorphs as for the normal organisms. This difference reflects the greater difference between the adult and juvenile versions of the polymorphs, which can affect even underlying biological properties. For example, one might suspect that butterflies and caterpillars do not necessarily eat the same kinds of things, and so their digestive systems might be different; their modes of locomotion are clearly different, so their structure and musculature could well be different; because of their different structures, their habitats and behaviors must be somewhat different. Thus, if one learns something about the digestion or muscles of caterpillars, one should be careful in attributing the same property to butterflies, even though it is seemingly a deep biological property. In short, subjects were appropriately acknowledging the dual nature of polymorphs, by hedging their bets in induction. This result parallels the similarity ratings of Study 1.

10. Study 4

This study tested children in an induction task using “blank” (meaningless) biological properties. One might criticize the more realistic properties of the previous study as reflecting specific knowledge of the categories; perhaps people already believed that border collies need to eat frequently, for example, and so were not truly inducing this property from the juvenile to adult dog. Although we attempted to choose properties lacking such specific knowledge (and we made up most of the biological properties), the use of blank properties eliminates this possibility. To this end, we tested adults as well as children, so that we can compare the results of Studies 3 and 4. As in Study 2, we used a forced-choice method so that children would not have to produce ratings. We expected that children would show a strong preference for induction to the adult version in the normal items; the main question was whether the same pattern would be found in polymorphs. To the degree that children are relying on surface similarity, their induction should weaken for such items. College students in Study 3 showed only a slight decline in inductive strength to the polymorph adults.

Method

Participants. Twenty-four children (11 male) from a local public second grade ($M_{age} = 7.0$), 26 older children (13 male) from local public fourth and fifth grade classrooms ($M_{age} = 9.8$) participated in this study. Sixteen New York University community members participated in an adult version of the task.

Materials and Procedure. Similar picture books using the same triplets as in previous experiments served as stimuli. Only internal biological properties were tested. Children viewed two labeled pictures and were told the labels of each item. The items included the adult form of an animal and the perceptual control used in
Table 5: Results of the induction question used in Study 4. Values close to 100 indicate choice of same species, values close to zero indicate choice of perceptual control item from other species.

<table>
<thead>
<tr>
<th>Age</th>
<th>Normal items</th>
<th>Polymorphs</th>
</tr>
</thead>
<tbody>
<tr>
<td>6- to 7-year-olds</td>
<td>69</td>
<td>48</td>
</tr>
<tr>
<td>9- to 10-year-olds</td>
<td>63</td>
<td>58</td>
</tr>
<tr>
<td>Adults</td>
<td>91</td>
<td>69</td>
</tr>
</tbody>
</table>

the previous studies. The facing page contained a target item consisting of the corresponding juvenile form. Children were asked whether the target item shared a biological property with the adult form or the perceptual control. For example, in one set of items, children viewed pictures of a fish and a frog. The experimenter told them, “This is a fish and this is a frog,” and then said, “This frog has TRIKA muscles. This fish has ANGIO muscles.” Children were then shown a tadpole and were asked, “Do you think this tadpole has TRIKA muscles like the frog or ANGIO muscles like the fish?” Order of presentation of the properties and items was counterbalanced across subjects. The properties included substances, organs, or structures internal to the organism, all labeled with novel names. After the completion of the eight picture sets, children also answered the same knowledge questions as in Study 2.

Following the initial induction questions, 11 of the children in the youngest age groups and 16 of the children in the oldest age group were asked why they had made their particular choices. Justifications were coded into the following categories: 1) same animal (e.g., “they are the same thing”), 2) parent-offspring or life cycle relations (e.g., “it’s a baby butterfly” or “the tadpole grows into a frog”), 3) deep behavioral/dispositional properties (e.g., “they are both nocturnal”), 4) same life stage (e.g., “they are both babies”), 5) physical appearance (e.g., “they look the same”), and 6) behavioral similarity (e.g., “they both move the same”).

Results

Children. The children’s performance in the induction task are presented in Table 5. As can be seen, children in each age group were generally more likely to choose different life form of the same species for the normal items than for the polymorphs ($M = 66\%$ vs. 53%). This was confirmed by a significant effect of animal type in a 2 (animal type: normal, polymorph) by 2 (age) mixed ANOVA ($F(1, 48) = 7.7, p<.01$). This difference was marginally greater for the younger than for the older children (interaction $F(1, 48) = 3.0, p = .09$), but in general participants appeared to choose almost equally between the different life forms of the polymorphs and the perceptual control item. When children’s knowledge of life stage changes was entered as a covariate in the analyses, the significant effect of animal type dropped out. Thus, the overall effect appears to be driven by knowledge of life stage changes.

In examining children’s explanations for their inductive choices it was clear that children were generally using principled reasons to explain their choices. Only two of the children (one from the youngest and one from the oldest age group) said that they had responded randomly or provided “don’t know” responses. One of the younger children stated that she referred to the word labels to inform her inductions for all of the items. These children’s responses were not examined further. The results of the six types of coded explanations for children’s inductive choices are provided in Table 6. The majority of the explanations for both normal items and polymorphs involved reference to belonging
Table 6  Average number of different types of explanations provided in the induction task, Study 4.

<table>
<thead>
<tr>
<th>Age</th>
<th>Normal Items</th>
<th>Polymorphs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal Items</td>
<td>Polymorphs</td>
</tr>
<tr>
<td></td>
<td>Same Kind</td>
<td>Parent-Offspring Relation</td>
</tr>
<tr>
<td></td>
<td>Deep</td>
<td></td>
</tr>
<tr>
<td>6- to 7-year-olds</td>
<td>1.64 0.55 0.00 0.46 0.36 0.36</td>
<td>1.06 0.55 0.00 0.27 1.09 0.18</td>
</tr>
<tr>
<td>9- to 10-year-olds</td>
<td>1.56 0.63 0.13 0.81 0.56 0.00</td>
<td>1.06 1.06 0.00 0.18 0.64 0.36</td>
</tr>
</tbody>
</table>

Note. There were four items per child of each type. However, children could give multiple reasons for an answer, and some very low-frequency categories are omitted, so the numbers in each row do not sum to 4.0.

to the same species or some mention of the relation between parents and offspring. Very few deep behavioral properties were provided. However, there were a substantial number of explanations referring to the same stage of the life cycle, especially among the older children for the normal items. References to physical appearance were generally not as common as references to the same animal or parent-offspring relations but were more likely to occur for the polymorphs.

Interestingly, few children used a single type of explanation for their inductions for the normal items (22%) or polymorphs (19%). Instead, 78% of the children used two or three different types of explanations to explain their inductive choices for the normal items and 70% of the children used two or three different types of explanations for the polymorphs. No children used four different explanation types for the four different normal items, but 3 children (11%) did provide four different explanation types for the four different polymorph items. These results suggest that the explanations are driven by the specific items and children’s knowledge of the animal rather than a single response strategy.

We performed a post-hoc analysis of the explanations to see if they were related to children’s choices in the induction task. We grouped together “deep” explanations referring to being the same kind or having deep properties in common (same kind, parent-offspring relation, deep behavioral properties, biological properties) and “superficial” explanations referring to various kinds of resemblance (same life stage, physical appearance, similar behavior). Obscure answers, “I don’t know,” word-matching strategies, and guesses did not fall into either class of response and hence were not included. This analysis revealed that children’s responses generally matched their explanations. When children chose the same species, they provided deep explanations 110 times and more superficial properties only 14 times. In contrast, when they chose the perceptual control, they referred to a deep property only 10 times and a superficial property 62 times. We did not perform a chi-square analysis to test this association, because the individual items are not independent (each child contributed up to 8 data points). Yet, it is obvious that there is a strong relation between children’s induction choices and explanations, as shown by a phi-coefficient (measuring association from 0-
This pattern was identical for the polymorph and the normal items, except that (as reported above), children chose the same species in the induction task more often for the normal items. But even for the polymorphs, when they chose the same species, they referred to deep properties much more than to superficial ones (50 vs. 6 times).

**Adult comparison.** Although Study 3 already tested adult subjects on these items, because the present properties and methodology (forced choice) differed from those used in the previous study, we thought it useful to test college students on the same induction task given to the children. The same pictures and properties were used, with age-appropriate instructions. As shown in Table 5, adults chose the property possessed by the same-species item more often for normal animals than for polymorphs ($t(15) = 3.42, p < .01$). That difference was much larger for adults than for children (22% vs. 13%). Adult subjects were also much more likely overall to choose the adult forms than the children were. Thus, adult subjects were influenced both by the developmental relationship between juvenile and adult organisms (in overall choice percentages) and perceptual-morphological similarity (in the difference between the item types).

**Discussion**

The main finding from Study 4 is that induction seems to be generally weaker from adult organisms to juveniles in polymorphs than in normal organisms. This study used only blank biological properties, which were designed to be shared across life forms, and these were not induced by children as strongly for the polymorphs as for the normal organisms. As in Study 3, this difference reflects the greater difference between the mature and juvenile versions of the polymorphs, which can affect even underlying biological properties. Thus, children, like adults, recognize the dual nature of these organisms. Note that children did not focus solely on perceptual resemblance, as their choices for the polymorphs were about evenly split across the perceptually similar control and the adult form. However, children’s explanations suggest that when they have background knowledge of the life cycle of different animals they are more likely to make inductions across dissimilar life forms of the same species than across animals that look alike and are at the same life stage. That is, with increasing knowledge (see next paragraph), children appear to be able to go beyond perceptual appearance in make inductions based on deeper biological relatedness.

Children provided a wide range of explanations for their inductions across different items. These results imply that the pattern of inductions and the explanations for particular inductions are strongly influenced by their knowledge of specific animal life forms. Of course, we cannot be sure of a causal link between the explanations and inference: Did children understand the “deep” relationship between parent and child and then use that as a basis for induction, or did they choose the kind of explanation that would support the induction they made? We suspect the former, because children had no prior knowledge about this property (e.g., angio muscles) and whether any item would have it. Therefore, they had to use the relation of the target item to the two base items in order to make a choice. It is revealing that children generally did not make counterfactual justifications for their superficial choices. For example, we only observed 10 occasions (out of 196 total explanations) in which they selected a perceptual control and then claimed that it shared a deep relationship with the target. If the explanations were constructed as after-the-fact justifications of the choice children made, more such responses would be expected.

Similar to the results of Study 2, children did not use the same type of explanation across items. One might have expected that children would form a response strategy and stick with
it throughout the questions, for example, always choosing the adult property or always relying on perceptual resemblance. Very few children gave the same explanation even within an item type (polymorph or normal). The variation in explanations shows that children were adapting their reasoning to the specific items and property being induced rather than carrying out a consistent strategy.

The results for adult subjects were consistent with the results of Study 3, using different properties and methodology. There we found that induction from juvenile to adult was stronger than from the juvenile to its perceptual control, for the deep properties. Here, using blank biological properties, we found that participants chose the adult form over the control the majority of time. Thus, the results of Study 3 cannot be due merely to prior knowledge involving the realistic features. Given that participants in this study had no prior knowledge suggesting that a butterfly would have trika or angio muscles, such an effect cannot account for the results.

11. General Discussion

Polymorphs are an interesting phenomenon for the psychology of concepts, because they test a number of underlying assumptions about how categories are constructed and used. Are they based on similarity or deeper properties? Do essences or appearances explain classification and induction? Our results suggest that simplistic answers to these questions are not likely to be correct. In fact, the results suggest that all the hypotheses have something right about them, and the problem for people is to negotiate these different bases of similarity and classification.

First, consider the issue of similarity. In the traditional classification literature, similarity has often been taken as a simple relationship between two stimuli indicating overall resemblance or feature overlap. However, others have pointed out that similarity has a habit of changing to fit the circumstances (Goodman, 1965; Medin, Goldstone, & Gentner, 1993; Murphy & Medin, 1985). If similarity is interpreted as measuring overlap of properties, then animals that are the same size, look similar, have similar behaviors, and so on, should be considered similar. Organisms differing in such properties but having a biological relationship should not be considered similar, as they have only one, nonsalient property in common.

In fact, our adult subjects clearly paid attention to both variables in making similarity and difference judgments. The most similar entities were those that were adult-juvenile pairs that also resembled one another. Pairs that were only perceptually alike or that only had a life-stage relationship were judged less similar. Thus, adults were clearly taking into account both the life-stage and perceptual information in judging similarity, and the two were about equally important in determining similarity.

Children who were asked to pick the item that was “most like” the target yielded a very similar pattern. They picked the adult version most often for the normal items, but they chose the adult version about half the time for the polymorphs (after the age of 5), suggesting that both life-stage information and perceptual resemblance went into their decision. It was only the youngest children who chose the perceptual control the majority of the time. However, this may not reflect an overall strategy of relying on perceptual similarity so much as ignorance, since only 50% of the 4- to 5-year-olds in Study 2 exhibited high levels of knowledge about life cycle changes, compared to 65% for 7- to 8-year-olds and 100% of children 9 and older.

In short, there is little evidence that any age group relied solely on perceptual variables or life-stage information in making their similarity judgments. Both sources of information seemed to influence judgments. Although it may not seem very surprising that people used both sources of information, this pattern is inconsistent with the views that people focus on underly-
ing biological properties or instead are phenomenologists in thinking about categories.

As we pointed out earlier, it is difficult for us to do the same comparison that Rips (1989) did between similarity and categorization, because the names we used refer to life stages as well as to species. That is, when referring to a member of the species \textit{Rana pipiens}, one cannot call it either \textit{tadpole} or \textit{frog} willy nilly. If it is a tadpole, to refer to it as \textit{my frog} is peculiar; if it is an adult, to call it \textit{my tadpole} is clearly wrong. Thus, unlike the novel names Rips made up or familiar names Hampton et al. used (e.g., bee vs. hummingbird), the names of familiar polymorphs do not lend themselves to asking whether classification changes across metamorphosis: The names refer to life stages rather than to the animal as a whole\(^1\). However, the existence of separate names in English suggests that metamorphosis was historically thought to create an important difference: A caterpillar changes into something else, which requires a distinct name. This is not the case with many animals in which the adult name is easily applied to the juvenile version (e.g., a kitten may be referred to as \textit{a cat} without anomaly — see footnote 1). Since linguistic classification does not provide a clear answer to whether people think of the juvenile and adult versions as the same kind of animal, we used induction as a measure of conceptual coherence. To the degree that different items share properties, they can be thought of as the same kind of thing. Furthermore, induction from one category to another has proven to be a good measure of conceptual closeness in a variety of materials (Klein & Murphy, 2002; Lin & Murphy, 2001; Rips, 1975).

The induction results showed that polymorphs did allow considerable induction (e.g., from caterpillar to butterfly). However, induction was stronger across the life stages of normal organisms (e.g., from border collie pup to border collie), again suggesting that both biological relatedness and perceptual similarity were important. And induction to perceptual controls (e.g., from tadpole to cod) was also sizeable — for example, receiving a rating of 55 (out of 100) even for biological properties.

By examining different predicates, we also were able to show that adult subjects had considerable inductive flexibility. When asked about deep behavioral and biological properties, they made inductions much more strongly from juveniles to adults than to the perceptual controls, but they reversed this preference for the superficial behavioral properties. This finding adds to past work with adults (Heit & Rubinstein, 1994; Ross & Murphy, 1999) and children (Kalish & Gelman, 1992; Nguyen & Murphy, 2003) showing that induction is not merely a function of the similarity between two instances or two categories. Rather, people evaluate the kind of relation between the two entities and whether it provides a basis for the particular property under consideration. A baby bunny and a baby mouse likely share properties related to their size, immaturity, blindness, and so on. A baby bunny and adult rabbit probably share properties based on physical and biological continuity. The shared size and immaturity allow for some kinds of inductions, and the biological continuity allows for

---

\(^1\) This difficulty is not completely symmetric, because in some cases the word for the adult form is polysemous, also referring to the organism as a whole. For example, the word \textit{frog} can sometimes be used as a covering term for the animal at whatever stage (e.g., \textit{Lots of frogs in this pond.}) or to refer to the adult form in particular (e.g., \textit{The tadpoles are in this aquarium, and the frogs are over there.}). The same is true of words like \textit{dog}, which can refer both to dogs in general and to the adult version in contrast to \textit{puppy}. However, there seems to be a greater infelicity in using the general term to refer to a specific juvenile when it is very different from the adult. That is, calling a tadpole a \textit{frog} or a caterpillar a \textit{moth} seems considerably odder than calling a puppy a \textit{dog}. However, for all of these cases, the term referring to the juvenile cannot be used generically to refer to juveniles and adults. Thus, both adult and juvenile terms refer to life stages, but the juvenile term lacks the additional generic sense. However, when in a task of naming the different life stages of an animal, it is presumably the life-stage sense of the word that people use, rather than the generic sense, causing the linguistic problem we have described.
others. These findings are a problem for any theory of induction that relies solely on similarity relations. Most authors have explained these effects as indicating reasoning from the relationship between the two items to a conclusion about whether the particular property would be likely to be shared (see Proffitt et al., 2000; Sloman, 1994).

For children (Study 4), we found that induction was stronger for normal items, even using biological properties, which would be the most likely to be common across life stages. But even children were willing to induce the property across life stages of polymorphs about half the time.

Taking the adults’ and children’s data together, one cannot state flatly that induction across metamorphosis is or is not very strong. Induction was stronger on average for normal items than for polymorphs, but induction also depends on the exact kind of property being induced, and adults had no difficulty in giving high ratings to inductions based only on biological relatedness (e.g., $M = 71$ for inducing biological properties across life stages of a polymorph) or based only on perceptual relatedness (e.g., $M = 70$ for inducing superficial behavioral properties to the perceptual control item). In short, if the property is the kind that is likely to be preserved across life stages, then people will induce it from one stage to another. Although that result might sound oxymoronic, it is not, because there might have been no features that people believed were likely to be preserved across life stages when metamorphosis intervenes.

**Metamorphosis and Essentialism**

As we remarked earlier, we do not believe that polymorphs can stand as the critical test case for essentialism. First, young children do not always know or understand the relationship of the two life stages. When someone tells them that the caterpillar “turns into” a butterfly, it is not very clear just what happens. Their experience with normal growth does not enable them to explain how a many-legged tubular creature turns into a winged creature with an entirely different shape and set of behaviors. Indeed, many adults (including one of us) are not really sure how this happens. At first, children may compartmentalize the information that “tadpoles turn into frogs” and treat this kind of change as quite different from changes that are part of the nature of the organism (Rosengren et al., 1997). For some children, the drastic changes associated with metamorphosis may be akin to a magic trick in which an egg is turned into a rabbit — it clearly has happened, but through a unique mechanism that is not a normal part of growth and development (Rosengren & Hickling, 2000; Rosengren, Kalish, Hickling, & Gelman, 1994). With education and continued experience, the child may come to realize that in fact metamorphosis is a natural change, even if it is drastic and not well understood.

Second, the fact that metamorphosis of this sort is part of the organism’s nature makes the question of whether an essence is preserved somewhat murky. An underlying essence is normally supposed to generate an entity’s superficial properties. In metamorphosis, most of those properties change, possibly suggesting a change in essence. But since this is a normal, programmed change, the change itself must reflect the creature’s underlying nature. A tadpole that never changes into a frog is not a proper tadpole. This case is quite different from the transformation experiments in which an external intervention is performed on a biological organism. And, indeed, Rips (1989) found that natural and artificial transformations engendered very different patterns of results in his classification and similarity judgments. Our findings suggest that adults at least are selective in how they apply essentialist reasoning to polymorphs. They can distinguish properties that should be induced from one life stage to the other from those that should not. Whether the animal changes its cat-
egory is perhaps a question that does not have a single good answer... it sort of changes its category, but sort of doesn’t. Our results show that people are aware of this ambiguity and judge similarity and induction accordingly. Indeed, perhaps Hampton et al.’s (2007) main finding is that there are significant individual differences in how people respond to the Rips scenarios. Our guess is that all their (adult) subjects understood the different ways that category membership seems to have both changed and stayed the same. They differed in which one they found the most important or salient.

Finally, although the notion of an essence has been very useful in explaining some aspects of categorization and reasoning about categories (Gelman, 2003; Medin & Ortony, 1989) as well as our own results, we should point out that essences did not dominate in our results: We found numerous effects of the inessential factor of perceptual similarity. For example, juvenile-adult similarity and induction were generally weaker for polymorphs than for normal entities that were perceptually similar. People chose the perceptual control fairly often (or rated it highly) in some conditions. On the other hand, purely perceptual accounts of induction (Sloutsky & Fisher, 2004) must acknowledge that knowledge of underlying biological relatedness supports induction, even when the entities are morphologically diverse and do not share the same name. Our results give concrete, repeated evidence that both biological relatedness and perceptual similarity were important in making judgments.

**Development**

We observed development from our youngest groups (4-5-year-olds) to college students across various tasks. In general, the youngest children showed the greatest difference between polymorphs and normal items. For example, they chose the same-species items least often by far in making most-like similarity judgments (Table 3). This is in contrast to Keil’s (1992) study, in which children of that age seemed to ignore differences between animals in developmental stages. Most likely the difference is due to task differences, as Keil used a conflict task pitting the animal’s end state against its developmental path, with unfamiliar animals.

At the other end of our age spectrum, college students often chose the same species items for polymorphs, and they gave high ratings to induction across life stages for polymorphs. However, adults did not globally give high ratings to polymorphs. Instead, they seemed to be choosing the precise form of similarity that was relevant to the particular question asked. As a result, adults favored the same species item for the most-like question much more often than children did, but they chose the same species item much less often than children did for the same-insides question. And in induction they gave low ratings to the adult polymorph for the superficial behavioral questions.

In short, adults seemed able to negotiate the conflicting forces of perceptual similarity and biological relatedness, focusing on one or the other depending on the task. Four-year-olds seemed less able to do so. Intermediate-aged children often gave responses somewhere in the middle of these two groups, both in classification and in induction (Table 5), suggesting that they were struggling with which aspect to emphasize. However, one should not interpret these results as reflecting an early global similarity stage followed by a use of deeper relations. First, the age results were heavily correlated with knowledge about the polymorphs, and when we introduced knowledge as a covariate, there was no main effect of age in either the similarity or induction tasks (Studies 2 and 4). When children were familiar with the transformations, they gave adult-like responses. Second, as we just discussed in some detail, adults did not rely solely on deeper relations, but used perceptual information in similarity judgments and induction both.

We believe that our results are consistent with
the general view that knowledge is driving development in this domain. Young children must first learn the facts about what the life stages of the polymorphs are. Older children may develop a better understanding of this change as a natural, programmed event. Finally, as adults, they are able to coordinate this knowledge with the particular judgment to be made. They can draw on different aspects of their knowledge based on their relevance to the question. Thus, rather than a global shift from relying on perceptual similarity to relying on biological properties, we observed a greater differentiation of how these items are both similar and different and a better understanding of when to use each of these important variables.

**Conclusion**

Rips (1989) concluded that people’s categorization judgments were not based on their judgments of similarity, arguing that an animal doesn’t change its category even if its external features change. Hampton et al. (2007), however, found that in most testing circumstances, animals that changed their form were perceived as then being more similar to and to be in the category of the later form — i.e., phenomenalism. Who is right?

Our suspicion is that both are right. When people are forced to make choices of this sort, there is no way for them to indicate nuanced understandings of the situation or even simple uncertainty. If an animal spontaneously transforms from a bird-like creature to a bee-like creature, then it is (eventually) similar to a bee, but it also may still have some properties of the bird-like thing it began with. Furthermore, contrary to the question posed in the Rips paradigm, it isn’t clear that it should be classified as *either* a bird or a bee, because neither animal undergoes a metamorphosis. Perhaps it is a new species altogether. But subjects are not able to make such responses in these experiments — they must choose one or the other category. Therefore, our focus on similarity and induction adds useful information to the classification data collected in previous studies in helping to form a more complete picture of people’s understanding of metamorphosis.

Our results show that both the perceptual changes and biological commonality are important to thinking about polymorphs, and they seem about equally salient in making similarity judgments. The contradictory results and individual differences found in the classification questions of Rips’s paradigm probably reflect slight changes in the emphasis of these two factors, causing subjects to weight one or the other more in that scenario/context/question format. Our results suggest that, with age, people seem increasingly able to think about both the continuity and the changes across the organism’s life stages and to coordinate them when thinking about polymorphs.

**References**


(Received 15 Sep. 2009)

(Accepted 30 Dec. 2009)

**Gregory L. Murphy**

Gregory L. Murphy received his PhD in 1982 from Stanford University in cognitive psychology. He taught at Brown University and the University of Illinois, and since 2001 he has been at New York University, in the Psychology Department. He has published widely in the field of the psychology of concepts, with special interest in hierarchies and category learning. In 2002, he published “The Big Book of Concepts” (MIT Press), which surveys the field of concepts. He served as editor of the Journal of Memory and Language and is currently associate editor of the Journal of Experimental Psychology: General.

**Karl S. Rosengren**

Karl S. Rosengren received his PhD from the University of Minnesota in 1989 and held a postdoctoral position at the University of Michigan. He has taught at Michigan and the University of Illinois. He is currently Professor of Psychology at Northwestern University, where he is also the Director of Undergraduate Studies. His research has focused on two areas of developmental psychology: cognitive development and motor control. He has a particular interest in children's understanding of magic.
### Appendix: Induction Features, Study 3

<table>
<thead>
<tr>
<th>Item</th>
<th>Biological</th>
<th>Deep Behavioral</th>
<th>Superficial Behavioral</th>
</tr>
</thead>
<tbody>
<tr>
<td>baby owl</td>
<td>has sesamoid bones</td>
<td>dies from exposure to below-freezing temperature</td>
<td>gives a warning cry when it sees a predator</td>
</tr>
<tr>
<td>border collie pup</td>
<td>has excellent color vision</td>
<td>needs to eat frequently</td>
<td>very active after dark</td>
</tr>
<tr>
<td>tadpole</td>
<td>requires biotin for hemoglobin synthesis</td>
<td>swims by moving its tail</td>
<td>eats mosquito larvae</td>
</tr>
<tr>
<td>baby bunny</td>
<td>has a strong immune system</td>
<td>leaves its nest very seldom</td>
<td>eats dandelion leaves</td>
</tr>
<tr>
<td>baby alligator</td>
<td>has an omentum in its abdominal cavity</td>
<td>eats mostly insects</td>
<td>uses its tail in fighting</td>
</tr>
<tr>
<td>monarch caterpillar</td>
<td>uses myoglobin for muscle contraction</td>
<td>moves around mostly by crawling</td>
<td>avoids direct sunlight</td>
</tr>
<tr>
<td>bean seed</td>
<td>has 12 chromosomes</td>
<td>will die without frequent rain</td>
<td>grows best in hot climates</td>
</tr>
<tr>
<td>bluebottle fly larva</td>
<td>manufactures hydrochloric acid during digestion</td>
<td>can swim</td>
<td>found near horses</td>
</tr>
</tbody>
</table>