Metamorphosis: Essence, appearance, and behavior in the categorization of natural kinds

JAMES A. HAMPTON
City University, London, England

AND

ZACHARY ESTES AND SABRINA SIMMONS
University of Warwick, Coventry, England

The transformation paradigm (Rips, 1989) was used to contrast causal homeostasis and strict essentialist beliefs about biological kinds. Participants read scenarios describing animals that changed their appearance and behavior through either accidental mutation or developmental maturation and then rated the animals on the basis of similarity, typicality, and category membership both before and after the change. Experiment 1 in the present study replicated the dissociation of typicality and categorization reported by Rips (1989) but also revealed systematic individual differences in categorization. With typicality and membership ratings collected between participants, however, Experiment 2 found no evidence for the dissociation and few essentialist responders. In Experiment 3, excluding information about offspring led most participants to categorize on the basis of appearance and behavior alone. However, with offspring information included and with questioning focused on the change of kind, essentialist categorization was still surprisingly rare. We conclude that strict essentialist categorization in the transformation task is relatively rare and highly task dependent, and that categorization is more commonly based on causal homeostasis.

In his famous story "The Metamorphosis," Franz Kafka describes a man who is transformed overnight into a gigantic insect. The story raises the intriguing question of whether the man has really become an insect, and if so, what (if anything) of his original identity has been retained. Modern science introduces a similar conundrum: If a tomato plant were genetically modified in such a way that it bore a fruit that exhibited the features of a grape, would that fruit still be a tomato, or would it have become a type of grape? And of course, children's fiction and folktales relate many instances of magical transformation (Kelly & Keil, 1985). Such metamorphoses—whether artificially induced or naturally occurring—provide a useful paradigm for investigating conceptual categorization. Dissociating an animal's appearance from its presumed biological essence allows for the observation of the relative influences of appearance and presumed essence on categorization. In the experiments reported here, we used a particular transformation paradigm, introduced by Rips (1989), to reexamine the evidence for essentialist categorization.

Rips (1989) used the transformation paradigm to demonstrate that categorization could be dissociated from similarity. The close relation often found between similarity and category judgments (Goldstone, 1994; Hampton, 1998, 2001; Rosch & Mervis, 1975) has resulted in several simil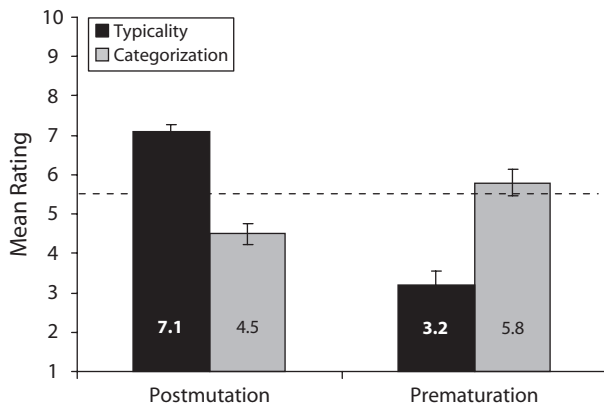arity-based models of categorization (see Murphy, 2002). According to these models, conceptual categories consist of clusters of similar objects, and instances are placed in the category whose contents they most closely resemble. Studies showing a dissociation between similarity and categorization are therefore of key interest, since they provide direct evidence against this view (Rips, 2001).

Rips (1989) generated two types of metamorphosis scenario. For ease of illustration, consider the example of a bird-like creature that came to look and act like an insect. In one condition (mutation), the metamorphosis was the result of accidental exposure to toxic waste. After its transformation, the creature nonetheless successfully mated with one of its original bird-like kind and had offspring that resembled its original kind (birds). Participants rated whether the transformed creature was more similar to, more typical of, and more likely to be a bird or an insect.

In the second condition (maturation), the story was of a creature with two naturally occurring life stages. Young bird-like creatures, called *sorps*, naturally matured into insect-like creatures called *doons*. When one insect-like doon mated with another doon, their offspring were bird-like sorps. Again, participants gave ratings, but in this case they concerned the sorp (i.e., the prematuration phase)—whether it was more similar to, more typical of, and more likely to be a bird or an insect.

J. A. Hampton, hampton@city.ac.uk

Rips (1989) reasoned that in the case of mutation, if people believe that a creature is what it is because of its deep essence and not because of its appearance, then they should see the animal in its final transformation as being more similar to and more typical of an insect; however, they should still rate the creature as more likely to be a bird. The mean ratings for typicality and categorization from the condition of Rips's study in which animals had undergone an accidental mutation are reproduced in the left panel of Figure 1. Because the ratings for typicality and similarity showed no differences, only typicality ratings are shown. In Figure 1, a low rating indicates that the initial category was chosen, and a high rating that the final category was chosen. Although the creatures were, after mutation, more typical of the final than the initial category (which should translate to higher ratings), categorization ratings were below the midpoint of the scale, implying that participants thought that the altered creature was more typical of an insect but more likely to be a bird.

For the maturation condition, Rips reasoned that belief in essences should cause people to see the adult form as the true category. In many similar cases of polymorphs in nature, we prefer to see the immature form as belonging to the same category as the reproductively active, adult form (a tadpole is a form of frog, not vice versa). This preference even extends to the female mayfly *Dolania americana*, which spends a year or more in aquatic nymph form but fewer than 5 min as a flying adult. The right panel of Figure 1 shows that typicality and categorization were again dissociated. Whereas participants judged the young sorp to be more typical of a bird than of an insect (lower ratings), they were nonetheless more likely to judge the sorp to be an insect—their categorization ratings were above

the midpoint. Rips thus obtained a neat double dissociation between type of scenario and the two kinds of rating.

Similar dissociations have been reported by others. Keil (1989) showed that when an animal's appearance was artificially transformed, older children (but not younger children) did not change their categorization of the animal. For instance, a horse painted with stripes to make it resemble a zebra was nonetheless categorized as a horse. Barton and Komatsu (1989) varied whether a given animal possessed the physical characteristics or the chromosomal structure of its kind. They found, for example, that a horse was no longer judged to be a horse when its chromosomal structure was changed (despite its physical similarity to horses) and that any animal with horse chromosomes was indeed categorized as a horse (despite physical dissimilarity to horses). Gelman and Wellman (1991) described an animal that had either its inner features (i.e., blood and bones) or its outward phenomenal features (i.e., fur and skin) removed; they then asked children whether the animal was still a member of its original category. As in the previous studies, Gelman and Wellman found that participants categorized according to the animal's deeper, or inner, features. Other dissociations of similarity and categorization have been demonstrated by Braisby (2004) and Thibaut, Dupont, and Anselme (2002).

The general interpretation of these results is that people do not categorize on the basis of appearance alone but, rather, believe that something deeper within the creature determines its kind. At this point, two different theoretical positions, which we term *causal homeostasis* and *essentialism*, concerning what this deep something may be can be distinguished. According to the causal homeostasis hypothesis (Boyd, 1991, 1999; Keil & Richardson, 1999), the mental representation of a natural kind concept includes causal links between different properties: A camel's hump enables it to go without water for long periods; a bat's hearing allows it to fly in the dark. When judging category membership, greater weight is placed on properties that are causes of other properties. Dissociation of typicality and categorization arises because typicality judgments place greater weight on appearance than on deeper properties—a typical example ought to look right as well as have the right stuff inside. This causal homeostasis view has been promoted by several researchers (Ahn, Kim, Lassaline, & Dennis, 2000; Hampton, 2001; Murphy & Medin, 1985; Rehder, 2003; Rehder & Hastie, 2001; Sloman, Love, & Ahn, 1998). The causal homeostasis hypothesis resembles two other theoretical positions. In many respects, it resembles what Gelman (2003) called *causal essentialism*,[1] the theory that people believe that all members of a kind share some deep property that causes their observable characteristics, that this property is the *causal essence* of the kind, but that possession of this essence may be a matter of degree. As Gelman and Hirschfeld (1999) put it,

subjects may believe that a certain inner quality or process of inheritance is needed in order for an animal to be a horse, but that in the real world different instances possess that quality or participate in that process to various degrees. (p. 409)

Figure 1. Typicality and categorization ratings for data estimated on the basis of Figures 1.5 and 1.6 in Rips (1989). We have reversed Rips's original scale, so that 1 = *initial category* and 10 = *final category*. The dashed line shows the midpoint on the scale. From "Similarity, Typicality, and Categorization," by L. J. Rips, in S. Vosniadou and A. Ortony (Eds.), 1989, *Similarity and Analogical Reasoning* (pp. 21-59), Cambridge: Cambridge University Press. Copyright 1989 by Cambridge University Press. Adapted with permission.

Causal homeostasis would suggest that instead of a single property that defines the essence of a kind, there may be a system of interconnected properties—functions and structures—that together play this role. Assignment of an individual to a category may then depend on the degree to which the deep properties of the individual (and their functional relations) match those found in a typical category member. Also related to the causal homeostasis position is Strevens's (2000) minimalist hypothesis, according to which people believe that some shared thing causes the members of a certain kind to have their observable characteristics but that there is no need to posit that this shared thing should be an essence. It is sufficient to suppose that people assume that "being of the kind" causes the characteristics, without specifying whether that shared thing is an essence.

In support of the causal homeostasis account, the typicality–categorization dissociation can be eliminated when participants are instructed to consider all information in their typicality judgments. The two measures, typicality and category membership, converge when equal weight is given to deeper causal features in both judgments. Goldstone (1994) reported an unpublished experiment using Rips's (1989) materials in which he elicited conceptual similarity ratings by asking questions such as "Which species is this animal more like, taking into consideration all of the information that you have available?" Given these instructions, similarity judgments were no longer dissociated from category judgments. Kalish (1995) also reported the dissociation of typicality and categorization in transformed natural kinds like those used by Rips (1989); however, when the judgments were made from a biological perspective (e.g., "Biologically speaking, how typical is . . . ?"), the dissociation was reduced or eliminated.[2] (See also Ahn & Dennis, 2001, for a similar result.) Thus, the dissociation of similarity/typicality and categorization depends critically on how typicality is understood.

The major alternative theoretical position to the causal homeostasis account is that people believe that natural kinds are categorized on the basis of an *essence* (Medin & Ortony, 1989; for an insightful review of notions of essence, see Gelman, 2003, pp. 3-14, and Rips, 2001; for an empirical study of essentialist beliefs in the social domain, see Haslam, Rothschild, & Ernst, 2000). An essence is more than a set of innards such as blood or bones. The naive notion of an essence for a biological kind can be assumed to involve several aspects (Gelman, 2003, pp. 10-11; Haslam, 1998; Rips, 2001; Strevens, 2000). The essence is something that an organism inherits from its parents and passes to its offspring; it is unchanging and indivisible; it is present in all members of the same type and only those members, resulting in sharp category boundaries; it is the primary cause of the common appearance, innards, and behavior of individuals of that type; and it justifies inductive projection of properties within the class. The essence is typically considered to be some intrinsic microstructure, hidden from casual inspection but deducible through scientific discovery. For example, many people may believe that genes, embodied in DNA, contain this essence for biological kinds and that this ge-

netic essence differentiates the species. Importantly, people may have little idea of what actually constitutes the essence, and it may thus act as an empty placeholder for the definition of the category (Keil, 2003; Medin & Ortony, 1989). The key difference between this notion of essence and the causal homeostasis view is that for an essentialist, neither deep nor surface similarity determines categorization. Characteristics of a creature (be they appearance, behavior, or deep causal processes) may be indicative of the presence of the essence, but they never constitute the reason for categorization.

The transformation studies introduced by Rips (1989) are particularly interesting and important for discriminating between these two accounts. On the one hand, the mutation in the creature implicates a change in deeper causal processes (unlike Keil's "zebra," with its painted stripes): The contamination enters the creature's body and works its changes from the inside. On the other hand, the unchanged offspring imply that the genetic essence has remained intact. Because the offspring have inherited the essence unchanged, the essentialist model predicts that the transformed creatures should still be of the same kind. An individual basing categorization on causal homeostasis, however, may easily suppose that the creature has changed category, given that the transformation produced deep changes in the causal structure of the organism. If the creature had acquired internal processes like those of the kind of creature that it resembled in the final phase, then the creature could be reclassified into that category, regardless of its offspring.[3]

The maturation scenarios are less directly relevant to this distinction. Natural dimorphism may be understood in various ways, and in fact there was considerable early debate in biology on how to handle such cases (Mayr, 1982; see also Murphy & Rosengren, 2006). According to the essentialist view, a creature's essence remains constant throughout its life cycle. However, this view is consistent with categorizing the creature according to either its younger or its older phase. Adult forms may be more revealing of the correct taxonomic class in real biology (which may have led Rips [1989] to predict that both forms would be categorized as belonging to the final phase), but since we are in the realm of fantasy biology here, people's essentialist intuitions may be open to either possibility.

Given the theoretical importance and wide influence of Rips's (1989) study (a recent search revealed over 180 citations), our aim was to explore the dissociation of typicality and categorization under more controlled conditions and to clarify the source of the effect. How robust is this much-cited phenomenon? Our motivation for exploring the effect came from inspection of the mean categorization ratings shown in Figure 1. Although the categorization and typicality means do indeed lie on opposite sides of the midpoint in both conditions, the two categorization means do not deviate far from that midpoint. Mid-scale categorization may reflect some combination of (1) consistent choice of mid-scale ratings, (2) consistent choice of extreme ratings by different individuals with roughly equal numbers choosing high or low, and (3) random responding across the scale. If (1) were true, then there

would have been a consensus that the creature was equally likely to be in either category. Such a finding would not support the interpretation that is now standard for these data. If (2) were true, then different groups of participants would have given systematically opposed ratings. Indeed, individual differences in this type of task have been reported elsewhere (Braisby, Franks, & Hampton, 1996; Malt, 1994; Smith, Patalano, & Jonides, 1998). If (3) were true, then the question simply would have been too vague to elicit any systematic responses, severely undermining the evidential validity of the paradigm.

EXPERIMENT 1

Experiment 1 was a conceptual replication of Rips's (1989) transformation experiment, with the following changes in materials and procedure. Rips collected similarity, typicality, and categorization ratings. Because similarity and typicality ratings were nearly identical in Rips's study, we collected only typicality and categorization ratings. Second, Rips gave the animals distinctly labeled stages of life in his maturation condition but only one label in his mutation condition. In our experiment, we did not use labels in either condition, thus making the conditions more directly comparable. Third, Rips explicitly asked participants to judge the animal only in its initial phase for the maturation condition and did not specify which phase should be judged in the mutation condition. Although the implication was that the final phase was the relevant one, the questions simply referred to the creature as, for example, "this doon (the one that changed)." To remove any possible ambiguity, in the present study, all participants read the full story about each creature, and on half the trials they rated the animal before it changed, whereas on the other half they rated it after it changed. Finally, our stories were more concise and direct. In order to shorten and simplify the task, we, unlike Rips (1989), made direct reference to the creature's resemblance to each category (see Appendix). Thus a creature initially "looked and acted just like a hummingbird" and eventually it "looked and acted just like a bee," whereas its offspring "looked and acted just like a hummingbird." This formulation meant that the similarity of a creature to a category could be taken to be equally high across all creatures and categories.

For Rips's (1989) pattern of results to obtain, we should find a three-way interaction of cause (mutation vs. maturation), phase (initial vs. final), and question (typicality vs. membership). For instance, before a hummingbird-like animal matures into a bee-like animal, it should be judged to be typical of a hummingbird but categorized as a bee. But after a hummingbird-like animal mutates into a bee-like animal, then it should be judged to be typical of bees but categorized as a hummingbird. The mutation scenarios were key to differentiations between causal homeostasis and essentialism. Essentialism predicts that people should not judge that mutation changes a creature's category, whereas causal homeostasis predicts that they may in fact do so (provided that the change in the creature's appearance and behavior is enough to infer a major restructuring of the causal processes inside).

Method

Participants. Thirty-two Princeton University undergraduates participated for partial course credit.

Materials and Design. We constructed 16 scenarios, each depicting an animal that comes to look and act like a different animal. All scenarios had the following standard format: (1) a brief description of an animal, including references to its looking and acting like an *x*; (2) a change in the animal due to either mutation or maturation; (3) a brief description of the changed animal, including reference to its looking and acting like a *y*; and (4) a statement that the animal's offspring resembled the initial state, *x*. Order of animals in each transformation was fully counterbalanced. No animal category appeared in more than one scenario. (See Appendix for a sample of the materials.)

Each scenario was followed by the two rating scales, which included space for participants to justify their ratings. Questions followed the format "[Before/After] it changed, was the animal more [typical of/likely to be] a hummingbird or a bee?" For each scenario, both questions (typicality, membership) asked about the animal in the same state (i.e., before or after the change), with the order of the questions counterbalanced across participants. Both scales ranged from 1, *initial category*, to 10, *final category*. Thus, we used a 2 (cause: mutation vs. maturation) × 2 (phase: initial vs. final) × 2 (question: typicality vs. membership) design, with all factors within participants. Booklets contained 16 scenarios, 1 per page, with 4 scenarios from each of the four conditions defined by the cause and phase factors. Order of scenarios was random for each booklet, and allocation of animal pairs to conditions was balanced across booklets.

Procedure. The participants were instructed to read each story carefully and then circle their answers on the rating scales. Following Rips (1989), the participants also listed the reasons for their decisions in the spaces provided below each story. These justifications served to increase the participants' engagement with the task and also provided evidence of how the participants reached their decisions.

Results

The ratings were submitted to an analysis of condition means and an analysis of individual differences in categorization patterns. Justifications were submitted to a content analysis and were related to the individual differences in categorization. These analyses are reported below.

Means analysis. Mean ratings are shown in Figure 2. Higher values indicate that participants rated an animal as more typical of, or more likely to be a member of, the final category. Comparing the middle two panels of Figure 2 with the results shown in Figure 1, it is clear that typicality and membership ratings were dissociated in the direction observed by Rips (1989), although in the present case the typicality means were more extreme, and the mean rating for prematuration categorization (3.01) fell below the midpoint. Two repeated measures ANOVAs were performed, one by participants (F_p) and one by items (F_i). For simplicity, Min F' is reported unless Min F' was marginally significant, in which case individual F_p and F_i are reported. Factors were cause (mutation, maturation), phase (initial, final) and question (typicality, membership). The three-way interaction was significant by participants [$F_p(1, 31) = 5.05, p = .03$] and marginal by items [$F_i(1,15) = 4.26, p = .057$]. Planned two-way ANOVAs conducted on the mutation and maturation scenarios separately showed highly significant interactions of phase and question [mutation: Min $F'(1,40) = 40.3, p < .001$; maturation: Min $F'(1, 39) = 28.8, p < .001$], thereby confirming the dissociation of typicality and categorization for each type of transformation.
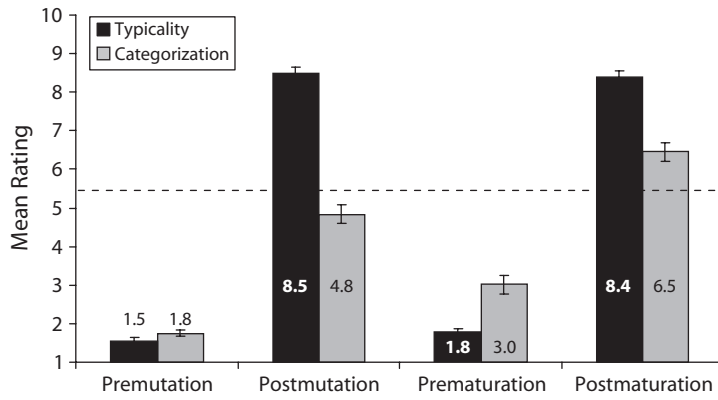
Figure 2. Typicality and categorization ratings from Experiment 1. Error bars indicate standard error of the mean. Ratings were made on a scale from 1, *initial category*, to 10, *final category*. The dashed line indicates the midpoint of the scale.

In Figure 2, the premutation condition confirmed that before their accidental contamination, animals were judged to be both typical of and likely to belong in the initial category. After the mutation, creatures were rated as clearly typical of their new category (8.5) but more likely to belong in their original category (4.8), just as in Rips (1989). In the maturation condition, typicality ratings also tracked the change in appearance (from 1.8 to 8.4), but categorization ratings changed less dramatically (from 3.0 to 6.5). The postmaturation condition was not measured in Rips (1989). In the prematuration condition, we failed to replicate his finding that in its young phase, the creature is more likely to be categorized according to the adult phase (mean rating of 3.0).

If participants held essentialist beliefs about categorization, then typicality ratings may have varied according to appearance across the different phases of an animal's life, but categorization ratings should have remained constant across both transformations. If an essence is taken to be an element of causal microstructure that is constant and inherited, and if a creature's category kind is determined by that essence, then a creature should never change category for any reason short of a change in the genetic information

encoded in its DNA. The data of Experiment 1 refute the hypothesis that this is how people think. Although categorization did not change as dramatically as typicality did, there were clear and significant shifts in categorization ratings accompanying each change. The fact that in its postmutation phase, an animal's appearance more closely matched that of a new category of animals clearly biased categorization ratings toward that category ($M = 4.84$ vs. $M = 1.75$). Given that the categorization ratings were again quite close to the midpoint of the scale (as in Rips, 1989), we next looked at the possibility that the means were obscuring systematic individual differences.

Individual analysis. Participants were classified on the basis of their choice of initial versus final categories for the categorization question for each of the four conditions. Using the rule that at least three of the four scenarios in each condition had to follow a prescribed pattern, we discovered that 26 of the 32 participants could be classified into just four groups. No more than 12 would be expected to meet the criterion if each question was judged independently of the rest, so responding was highly systematic [$\chi^2(1) = 26.1, p < .001$]. The remaining 6 participants dropped just below the criterion. Three of these were

Table 1
Proportions (Prop.) and Frequencies (Freq.) of Participants Exhibiting Each of the Defined Categorization Patterns in Each Experiment

| | Categorization Pattern | | | | | | | | Total Proportion of Consistent Patterns | |
| | Phenomenalist | | Rips Essentialist | | Origin Essentialist | | Nominal Essentialist | | | |
Experiment	Prop.	Freq.	Prop.	Freq.	Prop.	Freq.	Prop.	Freq.	Prop.	Freq.
1	.38	12	.16	5	.28	9	.09	3	.91	29
2 (Absolute)	.50	8	.12	2	.19	3	.00	0	.81	13
2 (Graded)	.69	11	.00	0	.12	2	.06	1	.87	14
3 (Standard)	.80	37	.00	0	.11	5	.02	1	.93	43
3 (Reduced)	.98	46	.00	0	.02	1	.00	0	1.00	47

Note—The four categorization patterns represented above are defined according to how participants classified each creature in each of four conditions: premutation, postmutation, prematuration, and postmaturation. The phenomenalist pattern is initial, final, initial, final, respectively, with "initial" indicating a judgment that the animal is a member of the initially described category and "final" indicating a judgment that it is a member of the finally described category. The Rips essentialist pattern is initial, initial, final, final; origin essentialist—initial, initial, initial, initial; nominal essentialist—initial, initial, initial, final.

easily and unambiguously classified into the pattern they most closely approximated, but the other 3 produced no consistent pattern (all 3 were essentialist for mutations). The corresponding patterns are shown in Table 1, together with the proportion and number of individuals showing each pattern.

The first major division of groups was based on the mutation scenarios. Twelve participants (38%) systematically judged that creatures changed categories (from the initial to the final phase) when the mutation took place, whereas the remaining 20 judged that they did not. All of the 12 also judged that maturation changed the creature's category, so we labeled these participants *phenomenalists*. They categorized in line with their typicality judgments, which reflected the outward appearance and behavior of the creatures. Regardless of the cause of the change or the nature of the offspring, these participants judged the creatures to be what they most closely resembled. In terms of the theoretical predictions, the phenomenalists showed a response pattern consistent with the causal homeostasis account. Since the creatures' entire appearances and behaviors had changed, participants now believed that their kind had also changed.

The remaining 20 participants were deemed to be essentialists, in the sense that they held that accidental mutation did not change a creature's category. However, we found that these participants categorized the maturation scenarios in three different ways. Five participants judged that a creature's pre- and postmaturation phases both belonged in the final category. We call these *Rips essentialists*, since these participants followed the pattern predicted by Rips (1989). A further 9 participants judged the opposite—that both phases belonged in the initial category. We termed this group *origin essentialists*, because they consistently categorized the creatures according to the appearance of the offspring. The final group, comprising only 3 participants, judged that, in the case of maturation (but not mutation), the category did change with the change in the creature's appearance; we call this group *nominal essentialists*. We hypothesized that these participants made their judgments on the basis of known cases of dimorphic species, such as the caterpillar/butterfly, and thus gave different names to the two forms. Finally, 3 participants gave responses in the maturation scenarios that fell exactly between two of the three patterns and were therefore left unclassified.

The discovery of these different response patterns illuminates the dissociations of typicality and categorization shown in Figure 2. In the premutation condition, all participants chose the initial category for both typicality and categorization. In the postmutation condition, mean categorization ratings were kept below the midpoint of the scale by the 20 essentialists who stayed with the initial category, whereas typicality ratings shifted to the final category for almost all of the participants, thus producing a reliable dissociation between the measures. At the same time, the 12 participants who judged that the category of the creature had changed led the mean categorization ratings to rise above the prematuration level. The mixture of different individual rating patterns also explains the dissociations observed in the maturation condition. Before an animal matures, only Rips's pattern of essentialism predicts that people will place the immature animal in the final category. Because only 5 of 32 participants clearly followed this pattern, the size of the dissociation in the prematuration phase was relatively small. In the postmaturation phase, the mean categorization rating was brought down toward the midpoint by the 9 origin essentialists who placed the creatures in the initial category.

Justifications. Participants were asked to write a brief explanation of their decisions for each scenario. Responses were classified into three major content areas. The first comprised appeals to appearance and behavior—"It looked and acted like a bee." These justifications were labeled *surface features* and were mentioned with high frequency across all conditions and groups. The second content area comprised appeals to the fact that the offspring were of a particular appearance. These we labeled *deep*. Finally, some participants appealed to the nature of the change that had occurred (i.e., mutation vs. maturation), and these we labeled *cause*. Figure 3 shows the proportion of justifications citing each type of content, broken down by the conditions and the three main patterns of responding. (There were not enough nominal essentialists for a meaningful discussion of their justifications to be given.)

Panel A of Figure 3 shows the justifications for phenomenalists, those who always said that the initial phase was in one category and the final phase in the other category, regardless of the cause of the change. As expected, all of these participants' justifications made reference almost every time to the appearance and behavior of the creatures. Examples from individual protocols justifying that mutation had changed categorization include: (1) "Ability to make a web is the key to an animal being a spider." (2) "The animal is closer to a bat because of its appearance and hunting method." (3) "[A] giraffe's behaviors depend on [its] long neck, therefore [the creature now is] more likely to be a camel."

Is it possible that these phenomenalist participants were not in fact basing their categorization on deeper causal principles at all but were simply following an appearance-based categorization strategy? The analysis of justifications suggests that this was not the case. Interestingly, about a third of the phenomenalists' justifications also referred to deeper features or offspring; for example, "It's more likely to be an octopus because of the description but [it] gives offspring that are like jellyfish so we lessen our confidence in the creature as an octopus."

Their justifications suggest that the phenomenalist participants reasoned that because both the behavior and the appearance had changed, there must be a new set of causal principles acting within the creature. Note the example in which the loss of the giraffe's neck was seen to have affected its ability to undertake its normal behaviors, causing the participant to judge that it had turned into a camel. The references to deeper features in the postmutation conditions shown in Figure 3A and the lack of references to the causes of the changes are consistent with this hypothesis. In their justifications, participants frequently expressed that they were impressed that not only a creature's ap-
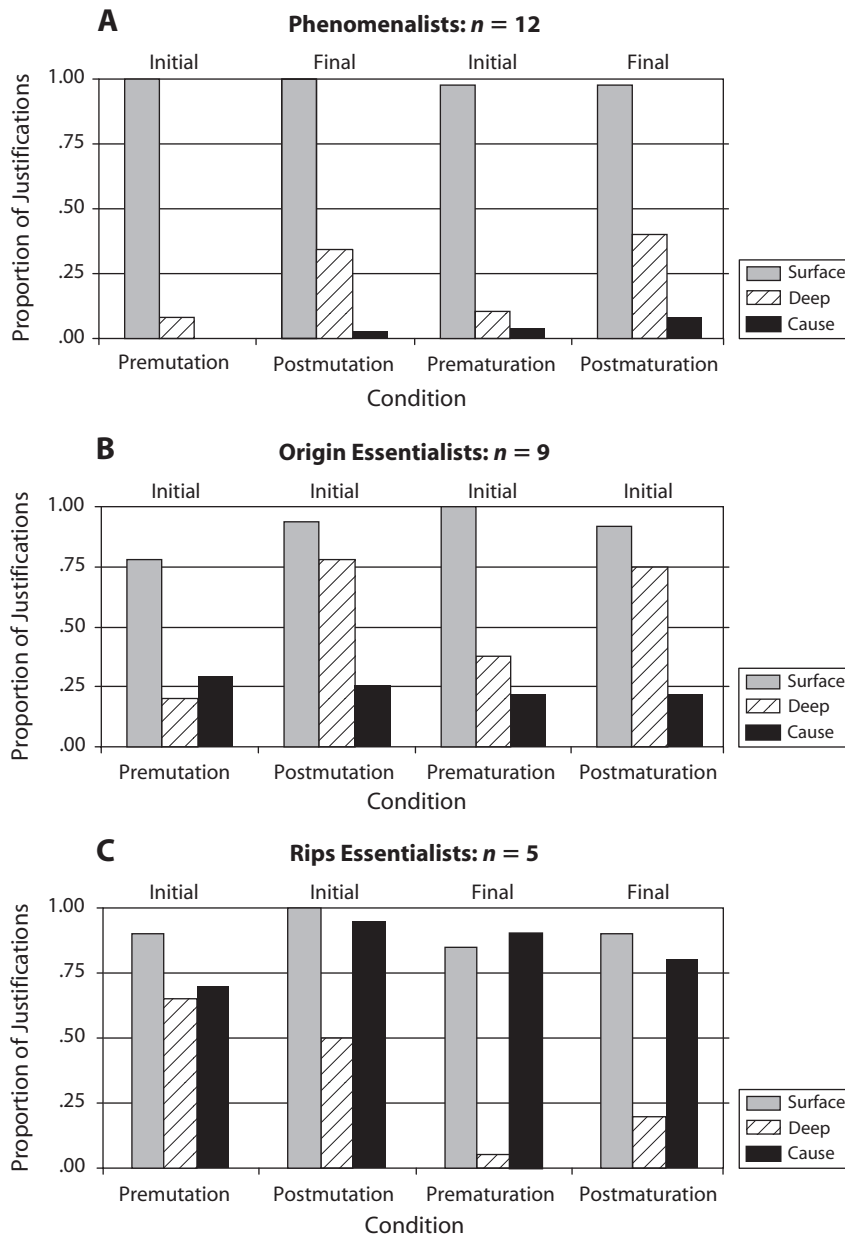
Figure 3. Proportions of justifications used by participants (classified into groups according to their categorization patterns) in Experiment 1. The labels *initial* and *final* above each set of bars refers to the cateory to which the creature in the scenario was assigned.

pearance but also its behavior had changed—presumably driven by changes in deep causal processes within the animal's nervous system. Closer inspection reveals that of the 12 phenomenalists, only 2 restricted themselves to justifications involving appearance and behavior alone; the other 10 mentioned deep or causal reasons for at least some scenarios.

The origin essentialists always placed creatures into the initial category, regardless of changes in appearance. As expected, their justifications regarding the final postmutation conditions included a high proportion of references to the offspring: (1) "Because [the] offspring were still wal-

ruses, genetics say it's a walrus." (2) "The animal was still a porcupine because it gave birth to porcupines." Participants in this group often made references to appearance and behavior but then discounted these features: "Though it looked and acted like a cat after the change, there was still part of the animal which told it to look and act like a squirrel."

In contrast, the Rips essentialists referred more frequently to the cause of the change. This makes sense, since this was the only group that differentiated between the two types of change (neither of the first two groups treated mutation and maturation scenarios differently). Appearance

and cause justifications were used with equal frequency, at around 80%. References to deep features, including to offspring, were found in the mutation conditions but not in the maturation conditions. Since it was in the mutation conditions that the Rips essentialists' choice of category coincided with the appearance of the offspring, it is clear that they were appealing to the offspring as evidence that no change occurred in the mutation scenario. Justifications for this group were clearly in line with similar justifications reported by Rips. In the mutation condition, they include: (1) "Since the offspring were beavers, it doesn't appear the changes were anything more than cosmetic." (2) "The change was induced by the pollutants and so can't change the species of the animal."

In the prematuration condition, this group assigned the creatures to the final category, offering justifications such as: "The beaver form is an earlier stage in the life of this porcupine-like animal." "It grew up to be a bat." "It's a frog–tadpole type situation."

Discussion

Experiment 1 not only replicated, under more strictly controlled experimental conditions, the dissociation of typicality and categorization ratings found by Rips (1989), but it also provided evidence that may explain the actual pattern of means Rips obtained. A notable aspect of Rips's data (and of our own) is the finding that unlike typicality ratings, mean categorization ratings were always close to the midpoint of the scale. Our analysis reveals that this effect is due to averaging across distinct groups of participants who exhibited very different patterns of categorization. A majority of participants were indeed essentialist categorizers in terms of the accidental mutation scenario, but only a minority of these followed the pattern proposed by Rips across both kinds of metamorphosis. Two other essentialist patterns were identified, and a significant minority (38%) of participants categorized on the basis of appearance and behavior alone, although their justifications strongly suggested that they were taking these observable changes as indicative of changes at a deeper level. The results therefore support both of the theoretical positions to some extent, with some participants adopting a causal homeostasis and some a strict essentialist strategy for answering the questions.

EXPERIMENT 2

Experiment 2 aimed to test the generality of the dissociation between typicality and categorization. In Experiment 1, participants gave two different judgments about each scenario. In this situation, participants might naturally have sought to differentiate the meaning of the two judgments, since they might have thought it could be considered uncooperative to work through the booklet giving identical answers to the questions about typicality and membership all the way through. The dissociation between the two scales might thus have been exaggerated in Experiment 1. If a within-participants manipulation induces people to contrast their typicality and membership judgments, then a between-participants design should eliminate the dissociation. Experiment 2 used the same materials as in Experiment 1, except that ratings (i.e., of typicality, membership) were collected between subjects. In addition, collecting the typicality and membership judgments separately made it possible to obtain separate justifications for each type of judgment.

Experiment 2 differed from Experiment 1 also in that it included another categorization condition. Rips's category membership judgment task asked for a decision about whether the animal was "more likely to be an x or a y." This question carries a presupposition that the creature definitely belonged in one or the other category. It may therefore have encouraged essentialist responding. To balance this, we introduced a second version of the category membership judgment (graded membership) that asked whether the animal was "more of an x or a y," implying that the creature's identity might have lain somewhere on a spectrum between x-ness and y-ness. Subtle differences of wording have had powerful effects in similar studies (see, e.g., Ahn & Dennis, 2001; Goldstone, 1994; Kalish, 1995). Thus, ratings of typicality and absolute membership (as in Experiment 1) and graded membership were collected from three different groups of participants.

Method

Participants. Forty-eight Princeton University undergraduates participated for course credit. None had participated in Experiment 1.

Materials and Design. The same materials as in Experiment 1 were used, with the exception that a graded membership question was added. We used a 2 (cause of transformation: mutation vs. maturation) × 2 (phase: initial vs. final) × 3 (question: typicality vs. absolute membership vs. graded membership) design with question as the only between-participants factor.

Procedure. The procedure was the same as that used in Experiment 1, with the exception that participants answered only one of the three questions for each scenario.

Results

Means analysis. It is clear from Figure 4, which shows mean ratings by condition, that membership judgments in both the mutation and maturation conditions closely tracked the typicality ratings in this experiment. That is, animals were almost always judged to belong to the category that they most resembled, leading to a postmutation shift in category membership. Most important, the cause × phase × question interaction observed in Experiment 1 was absent here. Participants appear to have categorized on the basis of appearance and behavior alone.

Factors in the ANOVA were again cause (mutation vs. maturation), phase (initial vs. final), and question (typicality vs. absolute membership vs. graded membership). The three-way interaction did not approach significance. Cause had a reliable main effect on ratings [Min $F'(1,46) = 4.3$, $p < .05$]. Participants rated the animals as being closer to the initial category when the cause of transformation was mutation than when it was maturation. The main effect of phase was also significant [Min $F'(1, 55) = 115.5$, $p < .001$], with low ratings for the animal in its initial phase and high ratings for the final phase. Neither the main effect of question nor any two-way interaction was reliable.
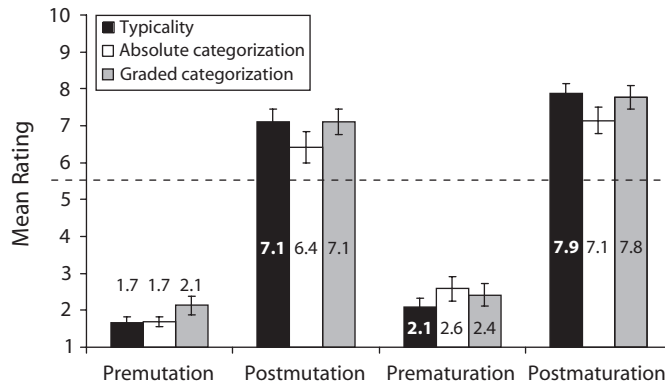
Figure 4. Typicality and categorization ratings from Experiment 2. Error bars indicate standard error of the mean. Ratings were made on a scale from 1, *initial category*, to 10, *final category*. The dashed line indicates the midpoint of the scale.

Individual analysis. Following the same analysis procedure used in Experiment 1, participants were classified into four groups. As Table 1 shows, in the categorization conditions, the majority of participants were phenomenalists. Essentialist responding, on the contrary, was rare. (Classification into patterns was also done for typicality, revealing 11 phenomenalists, 1 Rips essentialist, and 1 origin essentialist, very much in line with the patterns in the categorization conditions). Thus, phenomenalist categorization was the rule and essentialist responding the exception, with only 8 of 32 participants showing any of the three essentialist patterns.

Justifications. Given the relatively few participants generating consistently essentialist responses, a different way of displaying the relation between justifications and responding had to be adopted for Experiment 2. The two categorization conditions were combined and a count of the number of justifications of each of the three kinds (surface, deep, and cause of change) was made, broken down by the type of scenario (pre- vs. posttransformation and mutation vs. maturation) and the category selected (initial vs. final). Figure 5 shows the results of this cross-tabulation. Panel A shows (along the *x*-axis) that for mutation scenarios, before the transformation occurred, almost all of the responses (122) involved allocation to the initial category, with only 6 allocations to the final category (as a result, no justification data are reported for this condition). Of the 122 placements into the initial category, some 70% were justified using surface features, whereas deep and cause-of-change justifications were infrequent.

Following the mutation, justifications varied as a function of the category chosen. Those responses expressing the essentialist position that the creature was still in the initial category (*n* = 41) showed increased use of deep and cause-of-change justifications and reduced use of surface features justifications. In contrast, those who judged the creature to be in a different category after its transformation used justifications just like those used for categorization before the mutation.

Panel B of Figure 5 shows the data for the maturation scenarios. The sets of bars that represent decisions to place the young creatures into the initial category (*n* = 112) and the mature creature into the final category (*n* = 94) appear at the far left and far right of the figure, respectively; these bars show the same phenomenalist pattern of justifications mostly referring to appearance and behavior. Those participants who judged the young creature as belonging to the final category (*n* = 16) pointed primarily to the cause of the change as justification. Clearly, the fact that the process was one of natural development was of key importance to judging the immature form as properly belonging in the final category. By contrast, those participants who took the origin essentialist position and classified the adult form in the initial category (*n* = 34) seldom mentioned the cause of change as a justification; for these participants, the deep features (including offspring) justification was more prevalent.

Examples of justifications for different patterns of response include the following. Phenomenalist, postmutation: "It looked and acted like a snake but had lizard offspring." Origin essentialist, postmutation: "Toxic contamination just changed what it looked like, not what it actually was." Rips essentialist, postmaturation: "Since it's developmental, it must have always been a rabbit." Origin essentialist, postmaturation: "I still think it was a rhino—[it] had rhino offspring."

As in Experiment 1, the data were examined to see how many of the 16 participants in each of the three conditions restricted themselves to justifications referring only to appearance and/or behavior. In the typicality condition, there were 9 such participants; in the absolute membership condition, 4; and in the graded membership condition, 5. All of these participants showed the phenomenalist pattern. Therefore, about half of the phenomenalist categorizers (4 of 8 in the absolute membership condition and 5 of 11 in the graded membership condition) may have been truly phenomenalist, whereas the remainder were apparently also taking offspring or cause factors into account.

A   Mutation Scenarios

B   Maturation Scenarios
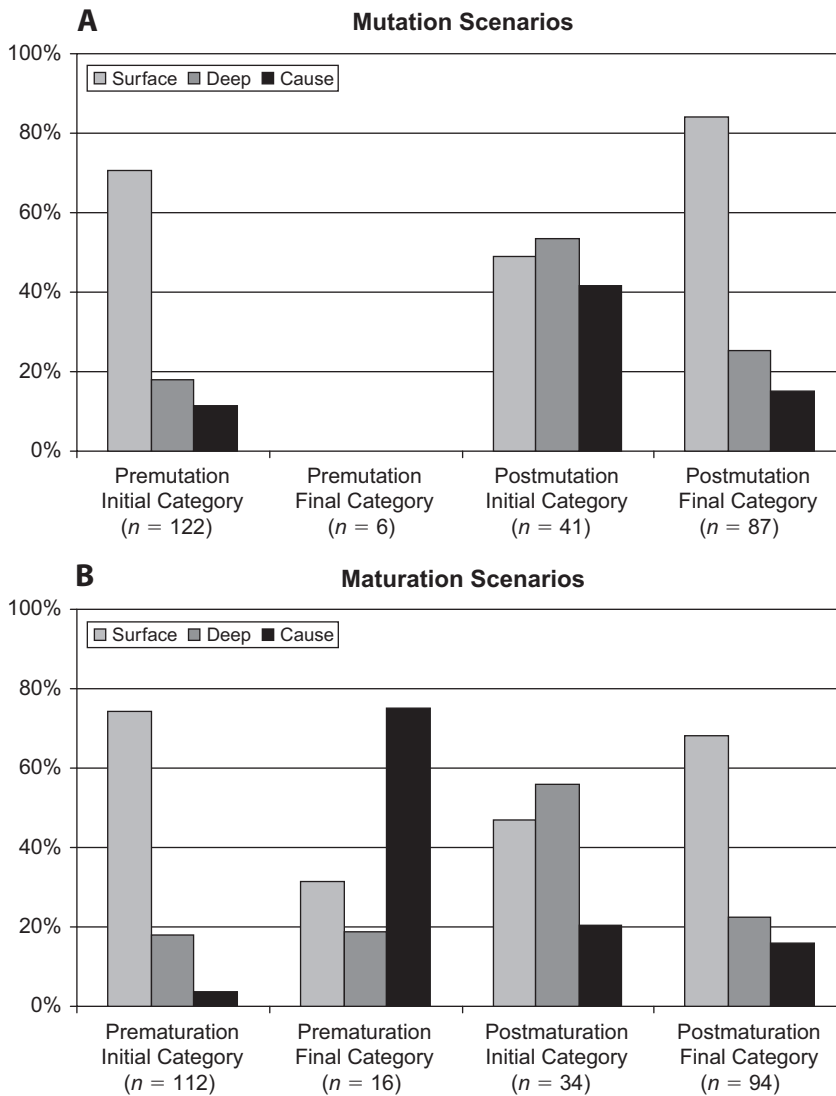
Figure 5. Percentages of justifications used in Experiment 2, broken down by scenario type and category chosen. Note—In the pretransformation category of the mutation scenario, data for the choice of "final category" were insufficient to plot.

In sum, the justifications given in Experiment 2 provide good support for the patterns of reasoning identified in Experiment 1. Classifying on the basis of appearance and behavior was largely justified in just those terms, for both kinds of change. Judging that a mutation did not alter an animal's category was justified through mention of offspring and the cause of the change. Judging that an immature form properly belongs in the adult category was justified by appeals to the cause of change, whereas the opposite judgment tended to appeal more to the appearance of the offspring.

Discussion

Experiment 2 failed to dissociate typicality and categorization ratings overall, and its results suggest that a siz-

able majority of participants in the categorization conditions categorized on the basis of appearance and behavior alone. The fact that typicality and categorization judgments were collected within participants in Experiment 1 does therefore appear to have encouraged essentialist responding. Individual analysis confirmed that people tended to adopt one of four patterns of responding. The analysis of the justifications revealed that for mutation scenarios, people sometimes focused on the fact that a creature's appearance and behavior had changed, whereas at other times, they focused on the fact that the creature's offspring had not changed. For maturation scenarios, animals were sometimes classified on the basis of appearance and behavior, sometimes both the immature and the mature animals were classified in the adult category on

the basis of the analogy with natural dimorphism, and at other times, a focus on the offspring led to both forms' being judged to belong in the initial category.

EXPERIMENT 3

Experiment 3 served two purposes. The primary purpose was to provide a more direct test of essentialist categorization. We hypothesized that if the task focused attention on the question of whether each animal had changed category, participants would then be reluctant to endorse such categorical transience and would be more inclined to give an essentialist response. We therefore instituted two major changes in the standard condition of Experiment 3. First, the participants judged the category membership of each animal both before and after its transformation. In the previous experiments, essentialist categorization was inferred from judgments across scenarios—that is, participants never categorized the same animal both before and after its transformation. Second, in order to further clarify the category judgments, we used a two-alternative forced-choice methodology. In the previous experiments, participants rated the animals on a category membership scale (i.e., from 0 to 10). However, scalar category ratings confound judgments that membership in a category is partial with judgments of confidence that membership is full (Estes, 2004). Thus in Experiment 3, participants provided binary category judgments and a separate confidence rating. Experiment 3 differed from the previous experiments in that it included similarity ratings. Although Rips (1989) found no difference between similarity and typicality, other studies have dissociated the two measures (Rips & Collins, 1993). We therefore included both measures in the present study.

A secondary purpose of Experiment 3 was to investigate a factor commonly held to affect essentialist categorization—namely, the animal's offspring (see, e.g., Gelman & Wellman, 1991; Pothos & Hahn, 2000; Rosengren, Gelman, Kalish, & McCormick, 1991; Springer, 1996). Given that offspring are considered to be indicative of an unchanged essence, removing information about offspring should generally reduce essentialist categorization. In addition to the standard condition described above, we therefore created the reduced condition, which was identical except that information about offspring was omitted. In the mutation scenarios of Experiments 1 and 2, an essentialist response could indicate one of two beliefs. One is the belief (based on the information about offspring) that a creature's germ line is unchanged, that genetics determine category kind, and that therefore the creature has not changed kind. The second possibility stems from a deeper essentialist belief—namely, that a creature can never change its kind through an external cause such as toxic contamination—it remains what it was born as, come what may. The reduced condition provided a test of these two alternatives. To the extent that essentialists respond to genetic information inferred from offspring, the reduced condition should make them more willing to entertain the possibility that the essence of the creature has also been changed by the mutation and thus respond by placing the

creature into the final category. To the extent that they are dyed-in-the-wool essentialists, they will continue to classify the mutated animal in its initial category even when no information is provided about offspring.

Thus, the experiment used a 2 (condition: standard vs. reduced) $\times$ 2 (cause: mutation vs. maturation) $\times$ 2 (phase: initial vs. final) mixed design, with condition a between-subjects factor and cause and phase within-subjects factors. The dependent measures collected from each participant were similarity, typicality, categorization, and confidence.

Method

Participants. Ninety-three undergraduates at the University of Georgia participated for partial course credit, 46 in the standard and 47 in the reduced condition. The participants were sampled in the same way, but those in the standard condition were run prior to those in the reduced condition. (We have no reason to suppose that the lack of complete randomization had any effect on our results.)

Materials. Stimulus materials were the same as those used in Experiment 1. For each scenario, participants provided all four judgments both before and after the animal's transformation. The order of the judgments—similarity, typicality, categorization, and confidence—was constant across scenarios and participants. Similarity and typicality were rated on a scale from 1 to 7, where 1 indicated the *initial* and 7 the *final* category. Category judgments were binary, with one choice indicating the initial category and the other indicating the final category. Participants also rated their confidence, on a scale from 1 to 5, in the preceding category judgment. Each scenario concluded with an open-ended query of the participant's justification for her or his judgments. Finally, whereas Experiments 1 and 2 used paper booklets, the data for Experiment 3 were collected on a computer, with scenarios displayed on a standard screen and responses collected via the keyboard.

Results

As in Experiment 1, group means and individual patterns of response were analyzed. Justifications provided little new information and therefore are not reported, in the interests of space.

Means analysis. Mean similarity and typicality ratings (on a scale from 1 to 7) were calculated for each condition. In every case, the pretransformation creatures were judged to be highly similar to and typical of the initial category (M between 1.31 and 1.43; $SE = 0.10$); posttransformation, they were judged to be highly similar to and typical of the final category (M between 6.33 and 6.59; $SE = 0.12$). Regardless of condition or cause of change, similarity and typicality judgments simply followed the description of the creatures' appearance and behavior. The effect of phase was highly significant for both measures in a three-way ANOVA, and no other effects or interactions approached significance.

The mean probability of participants' selecting the final category in each condition is illustrated in Figure 6, and mean confidence ratings are shown in Figure 7. In both conditions, choice of category also followed the appearance and behavior of the creature. Probability of selecting the final category was analyzed with a 2 (condition: standard vs. reduced) $\times$ 2 (cause: mutation vs. maturation) $\times$ 2 (phase: initial vs. final) ANOVA. The main effect of phase was significant [Min $F'(1,106) = 1245.7$, $p < .001$], as was the main effect of condition [Min $F'(1,105) = 4.56$,
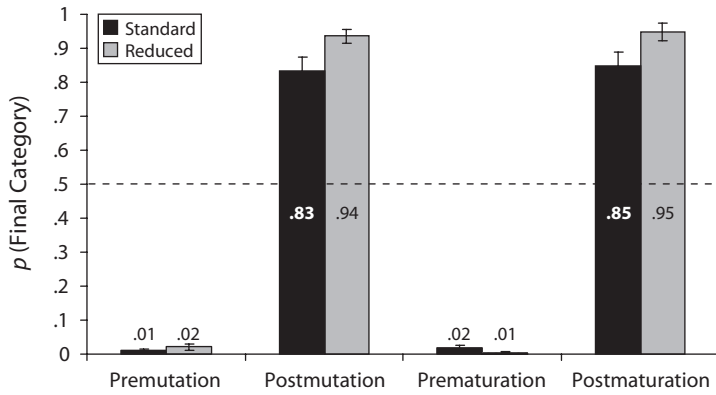
Figure 6. Mean categorization probabilities from Experiment 3. Error bars indicate standard error of the mean. The dashed line indicates the midpoint of the scale.

$p < .05$]. These two effects also interacted [$F_p(1,91) = 5.1, p < .05$; $F_i(1,15) = 42.6, p < .001$]. Before the transformation, the initial category was selected nearly 99% of the time. After the change, the final category was selected 84% of the time in the standard condition, rising to 95% of the time in the reduced condition in which offspring information was omitted.

Confidence ratings showed a significant effect of phase [Min $F'(1,106) = 7.8, p < .01$] and a significant phase × condition interaction [Min $F'(1,43) = 9.6, p < .005$]. As Figure 7 illustrates, before the change occurred, confidence was about level, at 4.4 for the standard condition and 4.3 for the reduced condition. After the change, participants in the reduced condition showed greater confidence ($M = 4.4$) than did those in the standard condition ($M = 4.0$). In the standard condition, the transformation induced some uncertainty about the animal's subsequent category, presumably because the offspring information was at odds with the appearance and behavior of the creature posttransformation. Note that none of these effects interacted with cause of change—the data were almost identical for the mutation and the maturation scenarios.

Individual analysis. Individual patterns of classification, scored using the same criteria outlined in Experiment 1, are reported in Table 1. (In contrast to Experiment 1, participants in Experiment 3 provided eight rather than four category judgments in each cell of the cause × phase design. The criterion for classifying a participant's responses now required that at least six of the eight responses in each of the four experimental cells follow the expected pattern.) Of the 46 participants in the standard condition, we classified 37 as phenomenalists, 5 as origin essentialists, and 1 as a nominal essentialist. None exhibited the Rips essentialist pattern, and only 1 (the nominal essentialist) responded differently as a function of the cause of change. The remaining 3 participants failed to fit any of the predefined response patterns. In summary, as the analysis of the means indicated, the vast majority of participants (80%) categorized on the basis of appearance, and only a small minority (13%) exhibited essentialist categorization.

The reduced condition was predicted to have less essentialist responding, and this was indeed the case. Of the 47 participants in this condition, all but 1 gave consistent
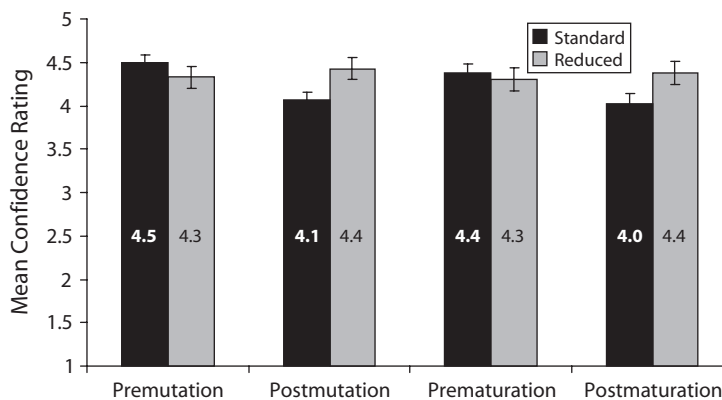
Figure 7. Mean confidence ratings from Experiment 3. Error bars indicate standard error of the mean.

phenomenalist responses, with the remaining individual being an origin essentialist. The increase in the proportion of phenomenalist participants from the standard to the reduced condition closely approached significance on a Fisher exact test ($p = .051$, two-tailed).

Discussion

Contrary to our expectations, very few participants in the standard condition endorsed essentialism. Prior to an animal's transformation, most participants judged it to be similar to, typical of, and likely to be a member of that initially described category, and subsequent to the transformation, it was judged similar to, typical of, and likely to be a member of the finally described category. Participants exhibited no reluctance to judge that the animal had changed categories as a result of its transformation, regardless of the cause of that transformation. As expected, when offspring information was omitted in the reduced condition, there was a further shift in the direction of phenomenalist response patterns. Placement of the changed creature into the final category rose, with increased confidence, and the number of essentialist participants dropped from 13% to 2%. We can conclude, therefore, that the majority of those who did respond in an essentialist way were strongly influenced by the evidence of genetic constancy in the offspring (i.e., they did not take the view that an external cause can never change an animal's kind).

The failure of the standard condition to increase essentialist responding was surprising. We expected that transparently asking about category membership both before and after the change would increase the amount of essentialist categorization. Instead, essentialist categorization actually decreased. It may be informative, therefore, to reconsider the differences between Experiments 1 and 3. Two differences are the inclusion of similarity ratings and confidence ratings in the present experiment. Neither of these changes seems a likely cause of the decrease in essentialism. Another difference is that in Experiment 3, the category judgments were within scenario and dichotomous, unlike in Experiment 1. Both of these changes were intended only to clarify the categorization task; we can think of no reason that either one should have reduced essentialist responses. Two final differences to consider are that the experiments were conducted on different student populations, and that whereas one experiment was done with paper and pencil, the other was computerized. Differences in participant motivation or understanding of the task may therefore also account for the different results.

What can be more firmly concluded is that essentialist responses are not easily replicated. In particular, the pattern of responding that Rips (1989) described as representative of his data was rarely seen in the present experiments. Of the 110 participants who categorized scenarios in Experiments 1 and 2 and in the standard condition of Experiment 3, 99 showed a consistent response pattern, but only 7 reproduced the classic pattern of essentialist categorization described by Rips.

Experiment 3 additionally revealed that in the standard condition, participants' confidence in their judgments declined after an animal had been transformed. This finding

makes intuitive sense in terms of the theory of causal homeostasis. Before transformation, there was no reason to doubt an animal's category, since it looked and acted one way and its offspring looked and acted the same way. But posttransformation, the animal's appearance and behavior contradicted those of its offspring. Since there was evidence both to support and to oppose placing the animal in the final category, participants' confidence was lower. In the reduced condition, in which the contradictory information had been removed, confidence was restored.

GENERAL DISCUSSION

The first experiment replicated Rips's (1989) seminal dissociation of typicality and categorization. Although a majority of participants (62%) did exhibit an essentialist pattern of categorization in relation to mutations—placing the creature in the same category before and after its change—mean categorization ratings still clearly changed across the transformation, as Rips (1989) also found. This change in the means was attributable to a substantial group of categorizers (38%) who remained resolutely phenomenalist, stating that the category of an animal changed when its appearance and behavior changed. When typicality and categorization judgments were separated between participants in Experiment 2, 59% of participants then categorized according to appearance and behavior, with only 25% showing essentialist patterns of response. When the question of change was put to participants more directly in the standard condition of Experiment 3 by asking participants to judge both the initial and the final phase of each creature, essentialist categorization was (surprisingly) again the exception rather than the rule. And as expected, removing the information about offspring (Experiment 3, reduced condition) reduced the incidence of essentialist categorization still further.

One clear conclusion is that essentialist categorization is highly dependent on the parameters of the task (see also Ahn & Dennis, 2001; Goldstone, 1994; Kalish, 1995). One could take Rips's (1989) study as an existence proof that at least under some conditions, some people will respond in an essentialist way. However, given the common citation of Rips's (1989) results as evidence for a more general view of how people understand natural kinds (Barton & Komatsu, 1989; Braisby, 2004; Gelman & Wellman, 1991; Hampton, 1995; Kalish, 1995; Keil, 1989; Pothos & Hahn, 2000; Rips, 2001; Sternberg, Chawarski, & Allbritton, 1998), our results severely limit the conclusions that should be drawn from Rips's study. Our results undermine any strong claim that people generally believe that accidental mutations do not affect categorization, even when the offspring are apparently unaffected.

In addition to these different patterns of response to the mutation transformation, we identified three different patterns in the maturation condition. Of those participants across all experiments who were essentialists for mutations, 7 consistently judged both phases to be in the final category, 19 judged both phases to be in the initial category, and 5 judged that the creature changed categories as it matured. (All of the phenomenalists agreed with the

last point of view for both mutations and maturations.) Heyman and Gelman (reported in Gelman, 2003, p. 287) also found large individual differences in the application of essentialist beliefs to traits such as intelligence, musical skill, or laziness. Similarly, C. L. Simmons and Hampton (2006) found strong correlations among different tasks reflecting essentialist beliefs, including whether membership is all or none or partial, whether one should defer to an expert about a categorization, and whether a categorization is a matter of fact or a matter of opinion.

What is the best account of the large number of participants who believed that the creature should be categorized according to its appearance? In the introduction, we identified two main hypotheses concerning people's beliefs about natural kinds. The causal homeostasis view holds that a creature is the kind it is because of deep biological processes that generate and maintain its appearance and behavior. Alternatively, the essence view holds that a creature's kind is determined by something like DNA or genetic essence, as observed in its offspring. When these two views are compared against our data, it is clear that the former view is about twice as common in the student populations sampled as the latter. The fact that posttransformation justifications often referred to deeper features in addition to appearance but rarely referred to the cause of the change provides evidence that changes in categorization were based on the belief that deeper causal processes had also changed. The increase in participants' confidence observed when offspring information was removed (Experiment 3) also suggests that the participants were considering information other than appearance and behavior but felt that the inferred deeper features outweighed the evidence of the offspring. Although the offspring presented good evidence that the germ line had not been affected, the change in deeper structure was sufficient to induce participants to change their view of the kind to which the creature belonged. The case of a person who undergoes sex reassignment surgery provides a parallel example: Although the person's genetic essence is presumably unaffected, friends and colleagues are asked to reclassify her or him in a different sex category. As the results from the participants in the present experiment indicate, there are likely to be strong individual differences in how easily this reclassification can be made.

The pattern of simple response data for the phenomenalists is also consistent with a different explanation—that these participants considered it most appropriate to name a thing on the basis of the conceptual category that it most closely resembled. When something looks and behaves just like a horse, then why not call it a horse, whatever its strange history may be? There is a tension between the notion of an ontological kind and the notion of a naming category (Sloman & Malt, 2003). For example, Malt (1994) showed that agreeing to call a liquid a kind of water relates not only to a belief that it contains H_2O but also to a belief about its origin and function. Naming or categorizing something on the basis of its appearance certainly makes good sense for pragmatic reasons—if, for example, one wanted to point it out to someone else

(Braisby, 2004). In folk taxonomy, there are many cases of species that share the same name although they are not in the same biological category; the Jerusalem oak and the poison oak, for example, are not in the oak genus *Quercus*. Thus, phenomenalist responders may have been reasoning pragmatically by categorizing the animals on the basis of similarity to other known animals. Understanding the categorization task as finding the best name for something would lead to changes in categorization following both mutation and maturation. It is possible that if people were instructed explicitly to contrast names for things against the underlying nature of things, a different pattern of responses would be found. Broadly speaking, we doubt that this explanation of our results is correct. We consider that a question such as "Is it more likely to be a squirrel or a rabbit?" would be understood as referring to ontological kinds rather than to naming categories. However, the issue deserves further exploration.

In sum, the essentialist pattern of categorization by which people maintain that a creature retains its kind through accidental or natural transformations by no means represents the only way that people react to stories of biological metamorphosis. In the majority of cases, we found that people were more likely to see a creature that had changed its appearance and behavior as having also changed its category, and we believe that this tendency will be general for studies of this kind. On the basis of the justifications given by participants, we argue that this change in category judgment is based on a causal homeostasis view of how natural kind categories are determined. People infer from the change in a creature's appearance and behavior that the deep causal principles of the creature now belong to a new category. This response pattern taken on its own, however, is also consistent with people deciding to categorize creatures on the basis of their appearance and behavior without any regard for internal causes or deeper knowledge. Whether the different patterns observed correspond to stable cognitive styles (e.g., essentialism, causal homeostasis, or nominalism) affecting other measures of conceptual processing, or whether individuals flexibly select among the various categorization strategies contingent on context and task factors, remains an important issue.

REFERENCES

Ahn, W. K., & Dennis, M. J. (2001). Dissociation between categorization and similarity judgment: Differential effect of causal status on feature weights. In U. Hahn & M. Ramscar (Eds.), *Similarity and categorization* (pp. 87-107). New York: Oxford University Press.

Ahn, W. K., Kim, N. S., Lassaline, M. E., & Dennis, M. J. (2000). Causal status as a determinant of feature centrality. *Cognitive Psychology*, **41**, 361-416.

BARTON, M. E., & KOMATSU, L. K. (1989). Defining features of natural kinds and artifacts. *Journal of Psycholinguistic Research*, **18**, 433-447.

BOYD, R. (1991). Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philosophical Studies*, **61**, 127-148.

BOYD, R. (1999). Homeostasis, species, and higher taxa. In R. A. Wilson (Ed.), *Species: New interdisciplinary essays* (pp. 141-186). Cambridge, MA: MIT Press.

BRAISBY, N. (2004). Similarity and categorization: Getting dissociations in perspective. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the Twenty-Sixth Annual Conference of the Cognitive Science Society* (pp. 150-155). Mahwah, NJ: Erlbaum.

BRAISBY, N., FRANKS, B., & HAMPTON, J. A. (1996). Essentialism, word use, and concepts. *Cognition*, **59**, 247-274.

ESTES, Z. (2004). Confidence and gradedness in semantic categorization: Definitely somewhat artifactual, maybe absolutely natural. *Psychonomic Bulletin & Review*, **11**, 1041-1047.

GELMAN, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford: Oxford University Press.

GELMAN, S. A., & HIRSCHFELD, L. A. (1999). How biological is essentialism? In S. Atran & D. L. Medin (Eds.), *Folk biology* (pp. 403-446). Cambridge, MA: MIT Press.

GELMAN, S. A., & WELLMAN, H. M. (1991). Insides and essence: Early understandings of the non-obvious. *Cognition*, **38**, 213-244.

GOLDSTONE, R. L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition*, **52**, 125-157.

HAMPTON, J. A. (1995). Testing the prototype theory of concepts. *Journal of Memory & Language*, **34**, 686-708.

HAMPTON, J. A. (1998). Similarity-based categorization and fuzziness of natural categories. *Cognition*, **65**, 137-165.

HAMPTON, J. A. (2001). The role of similarity in natural categorization. In M. Ramscar, U. Hahn, E. Cambouropolos, & H. Pain (Eds.), *Similarity and categorization* (pp. 13-28). Oxford: Oxford University Press.

HASLAM, N. O. (1998). Natural kinds, human kinds, and essentialism. *Social Research*, **65**, 291-314.

HASLAM, N. O., ROTHSCHILD, L., & ERNST, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology*, **39**, 113-127.

KALISH, C. W. (1995). Essentialism and graded membership in animal and artifact categories. *Memory & Cognition*, **23**, 335-353.

KEIL, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

KEIL, F. C. (2003). Categorisation, causation and the limits of understanding. In H. E. Moss & J. A. Hampton (Eds.), *Conceptual representation* (pp. 663-692). Hove, U.K.: Psychology Press.

KEIL, F. C., & RICHARDSON, D. C. (1999). Species, stuff, and patterns of causation. In R. A. Wilson (Ed.), *Species: New interdisciplinary essays* (pp. 263-282). Cambridge, MA: MIT Press.

KELLY, M. H., & KEIL, F. C. (1985). The more things change . . . : Metamorphoses and conceptual structure. *Cognitive Science*, **9**, 403-416.

MALT, B. C. (1994). Water is not H_2O. *Cognitive Psychology*, **27**, 41-70.

MAYR, E. (1982). *The growth of biological thought: Diversity, evolution, and inheritance*. Cambridge, MA: Harvard University Press.

MEDIN, D. L., & ORTONY, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179-195). Cambridge: Cambridge University Press.

MURPHY, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

MURPHY, G. L., & MEDIN, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, **92**, 289-316.

MURPHY, G. L., & ROSENGREN, K. S. (2006). *The two-body problem: Classification and reasoning about polymorphs*. Unpublished manuscript, New York University.

POTHOS, E. M., & HAHN, U. (2000). So concepts aren't definitions, but do they have necessary or sufficient features? *British Journal of Psychology*, **91**, 439-450.

REHDER, B. (2003). Categorization as causal reasoning. *Cognitive Science*, **27**, 709-748.

REHDER, B., & HASTIE, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General*, **130**, 323-360.

RIPS, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21-59). Cambridge: Cambridge University Press.

RIPS, L. J. (2001). Necessity and natural categories. *Psychological Bulletin*, **127**, 827-852.

RIPS, L. J., & COLLINS, A. (1993). Categories and resemblance. *Journal of Experimental Psychology: General*, **122**, 468-486.

ROSCH, E., & MERVIS, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, **7**, 573-605.

ROSENGREN, K. S., GELMAN, S. A., KALISH, C. W., & MCCORMICK, M. (1991). As time goes by: Children's early understanding of growth in animals. *Child Development*, **62**, 1302-1320.

SIMMONS, C. L., & HAMPTON, J. A. (2006, November). *Individual differences in essentialist beliefs*. Poster presented at the annual meeting of the Psychonomic Society, Houston.

SLOMAN, S. A., LOVE, B. C., & AHN, W. K. (1998). Feature centrality and conceptual coherence. *Cognitive Science*, **22**, 189-228.

SLOMAN, S. A., & MALT, B. C. (2003). Artifacts are not ascribed essences, nor are they treated as belonging to kinds. In H. E. Moss & J. A. Hampton (Eds.), *Conceptual representation* (pp. 563-582). Hove, U.K.: Psychology Press.

SMITH, E. E., PATALANO, A. L., & JONIDES, J. (1998). Alternative mechanisms of categorization. *Cognition*, **65**, 167-196.

SPRINGER, K. (1996). Young children's understanding of a biological basis for parent–offspring relations. *Child Development*, **67**, 2841-2856.

STERNBERG, R. J., CHAWARSKI, M. C., & ALLBRITTON, D. W. (1998). If you changed your name and appearance to those of Elvis Presley, who would you be? Historical features in categorization. *American Journal of Psychology*, **111**, 327-351.

STREVENS, M. (2000). The essentialist aspect of naive theories. *Cognition*, **74**, 149-175.

THIBAUT, J.-P., DUPONT, M., & ANSELME, P. (2002). Dissociations between categorization and similarity judgments as a result of learning feature distributions. *Memory & Cognition*, **30**, 647-656.

NOTES

1. Gelman proposes a second form of essentialism, *sortal essentialism*, that more closely resembles our understanding of the term *essentialism* in this paper. Theorists such as Gelman may see our two positions as being two versions of the psychological essentialist thesis. We have chosen to use the term *essentialism* in a manner more in keeping with its definition in social psychology as an unchanging, inherited, primary cause (Haslam, 1998).

2. The analysis was not reported in Kalish (1995), but we were able to confirm it using his data, which he kindly sent to us.

3. We do not assume that causal processes are more amenable to change in the fantasy world of the scenarios than are genetic essences. The point is that the radical metamorphosis of the creature may be taken as good evidence that the causal homeostasis system has changed to that of the new category, whereas the offspring's being of the original kind may be considered good evidence that the genetic essence has not changed.

APPENDIX

Four of the 16 scenarios used in Experiments 1–3 are described. The headings (e.g., *worm–snail*) were not included in the materials provided to participants. Wherever the element [*] appears below, the phrase "as a result of toxic contamination of its environment" in the mutation condition and "as a result of natural developmental processes" in the maturation condition appeared in the materials given to participant. In the reduced condition of Experiment 3, the final sentence was omitted.

1. Worm–snail: An animal had a segmented body with no arms or legs, and it burrowed into the soil sometimes. The animal looked and acted just like a worm. One day, [*], the animal began to change. It began to carry a small shell around on its back, grew two short antennae from its head, and left a slimy trail wherever it went. Finally, it looked and acted just like a snail. When it mated, the offspring looked and acted just like worms.

2. Snake–lizard: This animal had sharp front fangs, scaly skin, and a forked tongue. It looked and acted just like a snake. One day, [*], the animal began to change. It grew four legs, shed its fangs, and its tongue became sticky. The animal ended up looking and acting just like a lizard. After awhile it mated, and the offspring looked and acted just like snakes.

3. Hummingbird–bee: There was a small animal with wings and feathers, and it lived on the nectar of flowers. The animal looked and acted just like a hummingbird. But then, [*], the animal began to change. Eventually it ended up with transparent wings and a black and yellow striped body, always buzzing about. It looked and acted just like a bee. Then when it mated, the offspring looked and acted just like hummingbirds.

4. Horse–zebra: This brown animal had four legs, an elongated head, and a tail—it looked and acted just like a horse. But over time, [*], the animal began to change. It developed black and white stripes, and it came to look and act just like a zebra. When it mated, the offspring looked and acted just like horses.