Running Head: ESSENCE AND APPEARANCE

To appear in *Memory & Cognition, 2007*

Metamorphosis: Essence, Appearance and Behavior in the Categorization of Natural Kinds

James A. Hampton	Zachary Estes	Sabrina Simmons
City University, London	University of Warwick	University of Warwick

Please address correspondence to:

James A. Hampton

Psychology Dept, City University

Northampton Square, London EC1V OHB, UK

Email: hampton@city.ac.uk

Telephone: (44) 7040 8520

Abstract

The transformation paradigm (Rips, 1989) was used to contrast causal homeostasis and strict essentialist beliefs about biological kinds. Scenarios describing animals that changed their appearance and behavior through either accidental mutation or developmental maturation were rated for similarity, typicality and category membership both before and after the change. Experiment 1 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 by 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 categorization is more commonly based on causal homeostasis.

In his famous story *Metamorphosis*, Franz Kafka describes a man who is transformed overnight into a gigantic insect. This 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 so as to bear fruit that exhibit 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 folk tales have many instances of magical transformation (Kelly & Keil, 1985). Such metamorphoses—whether artificially induced or naturally occurring—provide a useful paradigm for investigating conceptual categorization. By dissociating an animal's appearance from its presumed biological essence, one may observe their relative influences on categorization. In the experiments reported below, we use a particular transformation paradigm, introduced by Rips (1989), to re-examine the evidence for essentialist categorization.

Rips (1989) used the transformation paradigm to demonstrate that categorization could be dissociated from similarity. The close relation normally seen between similarity and category judgments (Goldstone, 1994; Hampton, 1998, 2001; Rosch & Mervis, 1975) has resulted in several similarity-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 to which they have greatest similarity. Studies showing a dissociation between similarity and categorization are therefore of key interest as providing direct evidence against this view (Rips, 2001). Since Rips's (1989) paradigm served as the focus of our empirical investigation below, we next describe it in some detail.

Rips 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 change 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 also resembling 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 stages of its life. For example, young bird-like creatures, called "sorps", naturally mature into insect-like creatures called "doons". When the insect-like doon mates with another doon, they produce offspring which are bird-like sorps. Again, participants gave 3 ratings, but in this case they concerned the sorp (the pre-maturation phase). The sorp was rated for whether it was more similar to, more typical of, and more likely to be a bird or an insect.

Rips (1989) reasoned that in the case of a mutation, if people believe that a creature is what it is because of its "deep" essence and not because of its appearance, then participants should see the final phase of the animal's history as being more similar to an insect and more typical of an insect, but they should still rate it as more likely to be a bird. Mean ratings from his study are reproduced in Figure 1 (left panel). Rated typicality and similarity showed no differences so we present only typicality. In Figure 1 a low value indicates choice of the initial category and a high rating the final category. While creatures were more typical of the final category after mutation (higher values), categorization ratings were below the midpoint of the scale, implying that the altered creature was more typical of an insect yet 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 adult, reproductively active, phase (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 less than five minutes as a flying adult. The right panel of Figure 1 shows that typicality and categorization were again dissociated. Whereas the young sorp was judged more typical of a bird (low ratings), categorization was above the midpoint, that is, it was *more likely to be* an insect. 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 so as to 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 either had its inner features (i.e., blood and bones) or its outward phenomenal features (i.e., fur and skin) removed, and then they 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 by the "deeper" or inner features. Other dissociations of similarity and categorization have been demonstrated by Braisby (2004) and by 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 can be distinguished concerning what this "deep" something may be, which we shall term "causal homeostasis" and "essentialism". By 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 having the right stuff inside. This 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 bears similarity to two other related theoretical positions. In many respects it resembles what Gelman (2003) called "causal essentialism".[1] Her notion is that people believe that all members of a kind share some deep property which causes their observable characteristics, this property being the causal essence of the kind. In addition possession of this essence may be a *matter of degree.* As Gelman and Hirschfeld (1999, p. 409) 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".

Causal homeostasis would suggest that instead of a single property there may be a system of interconnected properties – functions and structures – which together play this role. Membership of an individual in a category may then reflect 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 minimalist hypothesis (Strevens, 2000) according to which people believe that something about being of a certain kind causes the members of that kind to have their observable characteristics, but that there is no need to posit that this "something" should be an essence. It is sufficient to suppose that people assume that "being of the kind" causes the characteristics, without specifying whether that something 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 converge when equal weight is given to deeper causal features in both tasks. Goldstone (1994) reported an unpublished experiment using Rips's (1989) materials, in which he elicited *conceptual* similarity ratings (i.e., "Which species is this animal more like, taking into consideration all of the information that you have available?"). Given these instructions, similarity was no longer dissociated from category judgments. Kalish (1995) also dissociated typicality and categorization in transformed natural kinds like those used by Rips (1989). But 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 and 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 this 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 naïve 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). It is something that an organism inherits from its parents and passes to its offspring; it is unchanging and indivisible; it is present in all and only members of the same type 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 genetic essence differentiates the species. Importantly, people may have little idea what the essence may be, so that it may act as an empty place-holder 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 will determine 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 major change 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 produces deep changes in the causal structure of the organism. If the creature had acquired internal processes like those of the final phase category, then the creature could be reclassified in that category, regardless of the 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). By the essentialist view, the creature's essence remains constant throughout the 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 classed with the final phase). But as 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 effect under more controlled conditions and to clarify the source of the dissociation. 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. According to (1), there is a consensus that the creature is equally likely to be in either category. Such a finding would not support the interpretation that is now standard for these data. According to (2) there are different groups of participants with 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). According to (3) the question is simply too vague to elicit any systematic responses. If this result were found, it would severely undermine 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, in Experiment 1 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 used labels in *neither* 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, while 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, unlike Rips (1989) we 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," while its offspring "looked and acted just like a hummingbird." This formulation meant that similarity of the creatures to each category could be taken to be equally high.

For Rips's (1989) pattern of results to obtain, we should find a three-way interaction of Cause (mutation, maturation), Phase (initial, final) and Question (typicality, membership). For instance, *before* a hummingbird-like animal *matures* into a bee-like animal, it should be judged 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 typical of bees but categorized as a hummingbird. The mutation scenarios were the key conditions differentiating between causal homeostasis and essentialism. Essentialism predicts that the mutated creature should not change category, whereas causal homeostasis predicts that it 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 reference to looking and acting like an X, (2) a change of the animal due to either mutation or maturation, (3) a brief description of the changed animal, including reference to 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.)

After each scenario were the two rating scales, and 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 order of questions counterbalanced across participants. Both scales ranged from 1 (initial category) to

10 (final category). Thus, the design was 2 (Cause: mutation, maturation) X 2 (Phase: initial, final) X 2 (Question: typicality, membership), all factors within-participants. Booklets contained 16 scenarios, one per page, with four 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. Participants carefully read each story, and then circled their answers on the rating scales. Following Rips (1989), participants also listed the reasons for their decisions in the spaces provided below each story. These justifications served to increase participants' engagement with the task and also provided evidence of how 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 related to the individual differences in categorization. These analyses are reported below.

Means analysis. Mean ratings are shown in Figure 2. Higher values corresponded to rating 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 pre-maturation categorization (3.01) fell below the midpoint. Two repeated measures ANOVA 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 will be reported. Factors were Cause (mutation, maturation), Phase (initial, final) and Question (typicality, membership). The 3-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.

In Figure 2 the pre-mutation condition confirmed that before their accidental contamination animals were judged 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). For maturation, typicality also tracked the change in appearance (from 1.8 to 8.4), but categorization changed less dramatically (from 3.0 to 6.5). The post-maturation condition was not measured in Rips (1989). For pre-maturation 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 hold essentialist beliefs about categorization, then while typicality ratings may vary with appearance across the different phases of an animal's life, categorization ratings should be 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. While categorization did not change as dramatically as did typicality, there were clear and significant shifts in categorization ratings accompanying each change. The fact that post-mutation an animal had the appearance of a new category clearly biased categorization ratings towards that category (M = 4.84 vs. 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 3 out of the 4 scenarios in each condition must follow a prescribed pattern, we discovered that 26 of the 32 participants could be classed into just 4 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, χ^2 (1) = 26.1, p < .001. The remaining 6 participants dropped just below the criterion. Three were easily and unambiguously classified into the pattern they most closely approximated, but 3 others produced no consistent pattern (all three 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 category (from initial to final) when the mutation took place, while 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 creature. Regardless of the cause of the change, or of the nature of the offspring, these participants judged the creature to be what it most closely resembled. In terms of the theoretical predictions, the phenomenalists were showing a response pattern consistent with the causal homeostasis account. Since the creature's whole appearance and behavior had changed, they now believed that its 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 the creature's category. However these participants were discovered to categorize the maturation scenarios in three different ways. Five participants judged that both phases of the creature belonged in the final category. We call these “Rips essentialists”, as this was the pattern predicted by Rips (1989). A further 9 participants judged the opposite – that both phases remained in the *initial* category. We termed this group “Origin essentialists”. Note that this group always placed a creature in the category corresponding to the appearance of the offspring. The third pattern, which we termed “Nominal essentialists” was observed in just 3 participants, and involved judging that in the case of maturation (but not mutation) the category *did* change with the appearance. We hypothesized that these participants were judging on the basis of the analogy with known cases of dimorphic species such as the caterpillar/butterfly, and so gave different names to the two forms. Finally, 3 participants gave responses in the maturation scenarios that fell exactly between two of the 3 patterns, and so were left unclassified.

The discovery of these different response patterns illuminates the dissociations of typicality and categorization shown in Figure 2. In the pre-mutation condition, all participants chose the initial category for both typicality and categorization. In the post-mutation condition, the 20 essentialists stayed with the initial category, whereas almost everyone’s typicality ratings shifted to the final category, thus producing a reliable dissociation between the measures. At the same time, because 12 participants judged that the category had changed, mutation increased mean categorization ratings. 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 categorization in the final category. Because only 5 of 32 participants clearly followed this pattern, the size of the dissociation pre-maturation was relatively small. After the animal matures, the mean rating was brought down towards the midpoint by the 9 Origin essentialists giving initial category judgments.

Justifications. Each participant was asked to write a brief explanation of the reasons for their decisions for each scenario. Responses were classified into three major content areas. First there were 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. Second there were appeals to the fact that the offspring were of a particular appearance. These we labeled "deep". Finally some appealed to the nature of the change that had occurred (i.e. mutation versus 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 insufficient Nominal essentialists for a meaningful discussion of their justifications to be given).

Taking the first panel, the figure for Phenomenalists shows the justifications for those who always said the initial phase was in one category and the final phase in the other category, regardless of the cause of the change (this pattern is labeled at the top of the graph). Naturally enough, all made reference almost every time to the appearance and behavior of the creatures. Examples from individual protocols justifying that mutation had changed categorization were:

"Ability to make a web is the key to an animal being a spider"

"The animal is closer to a bat because of its appearance and hunting method"

"[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 simply following an appearance-based categorization strategy? The analysis of justifications suggested that this was not the case. Interestingly, about a third of their justifications also referred to deeper features or offspring. An example was

> “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”.

Justifications suggested that 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 where the loss of its neck was seen to affect the giraffe’s ability to undertake its normal behaviors, so the creature was deemed to have changed into a camel. The reference to deeper features in the post-change conditions in Figure 3, and the lack of reference to the cause of the change are consistent with this hypothesis. Participants in their justifications were frequently impressed by the fact that the creature’s behavior had also changed –presumably driven by changes in deep causal processes within the animal’s nervous system. Closer inspection revealed that of the 12 Phenomenalists only 2 restricted themselves entirely to justifications involving appearance and behavior alone, the other 10 mentioning deep or causal reasons for at least some scenarios.

The Origin essentialists always categorized creatures in the initial category, regardless of changes in appearance. As expected, for the final post-change conditions there was a high proportion of references to the offspring. Examples were

> “Because [the] offspring were still walruses, genetics say it's a walrus”
>
> “The animal was still a porcupine because it gave birth to porcupines”.

This group often made reference to appearance and behavior, but then discounted it, as in

> “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 behavior makes good sense since this was the main group who differentiated between the two types of change (neither of the first two groups treated mutation and maturation scenarios differently). Appearance and Cause justifications were about equally frequent at around 80%. Deep features, including reference to offspring, were found in the mutation conditions, but not in the maturation conditions. Since it was in the mutation conditions that their choice of category coincided with the appearance of the offspring, it is clear that the Rips essentialists were appealing to the offspring as evidence for there being no change in the mutation scenario. Sample justifications for this group were clearly in line with similar justifications reported by Rips. For example, for mutation:

> "Since the offspring were beavers, it doesn't appear the changes were anything more than cosmetic"
>
> "The change was induced by the pollutants and so can't change the species of the animal"

For the choice of final category for pre-maturation, this group offered 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 Rips's (1989) dissociation of typicality and categorization ratings under more strictly controlled experimental conditions, but also provided evidence that may explain the actual pattern of means he obtained. A notable aspect of Rips's data (as well as our own) was the finding that unlike typicality ratings, mean categorization ratings were always close to the midpoint of the scale. Our analysis reveals that this effect was 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 the justifications strongly suggested that they were taking these observable changes as indicative of changes at a deeper level as well. The results therefore supported both of the theoretical positions to some extent, with different groups of participants adopting a causal homeostasis or 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 may naturally seek to differentiate the meaning of the two judgments. 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 may 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 Experiment 1, but ratings (i.e., typicality, membership) were collected between-subjects. In addition, by collecting the typicality and membership judgments separately, it was possible to obtain separate justifications for each type of judgment.

The other difference between Experiments 1 and 2 was the inclusion of another categorization condition. Rips's category membership judgment asked for a decision of whether the animal is "more likely to be an X or a Y". This question carries a presupposition that the creature is definitely in one or the other category. It may therefore encourage essentialist responding. To balance this, we introduced a second category membership judgment ("graded membership) that asked whether the animal is "more of an X or a Y", in which it is implied instead that the instance may lie somewhere on a gradient from Xness to Yness. Subtle differences of wording have had powerful effects in similar studies (e.g., Ahn & Dennis, 2001; Goldstone, 1994; Kalish, 1995). Thus, ratings of typicality, 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 were used as in Experiment 1, with the exception that a graded membership question was added. The design was 2 (Cause of transformation: mutation, maturation) X 2 (Phase: initial, final) X 3 (Question: typicality, absolute membership, graded membership), with Question as the only between-participants factor.

Procedure. The procedure was the same as for 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 both category membership judgments 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 post-transformation shift in category membership. Most importantly, the Cause x Phase x Question interaction observed in Experiment 1 was absent here. Participants appear to have categorized on the basis of appearance and behavior alone.

ANOVA factors were again Cause (mutation, maturation), Phase (initial, final) and Question (typicality, absolute membership, graded membership). The 3-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 closer to the initial category if the cause of transformation was mutation than if 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 2-way interaction was reliable.

Individual Analysis. Following the same analysis procedure as in Experiment 1, participants were classified into four groups. As shown in Table 1, the majority of participants in the categorization conditions were phenomenalist. Essentialist responding, on the contrary, was rare. (Classification into patterns was also done for Typicality, and showed 11 Phenomenalists, 1 Rips essentialist and 1 Origin essentialist, very much in line with 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 was made of the number of justifications of each of the three kinds (surface, deep, and cause of change), broken down by the type of scenario (before vs. after and mutation vs. maturation) and the category selected (initial or final). Figure 5 shows the results of this cross-tabulation. The top panel shows (along the x–axis) that for mutation scenarios, before the change occurred, almost all responses (122) were for the initial category, with only 6 for the final category (as a result no justification data are reported for this condition). For these 122 categorizations in the initial category, some 70% were justified using surface features, while deep and cause of change justifications were infrequent.

Following the mutation, justifications varied as a function of the category chosen. Those responses taking the essentialist position (N = 41) that the creature was still in the initial category showed increased use of deep and cause of change justifications, and reduced use of surface features. By contrast those judging the creature now to be in a different category show a pattern of justification just like that for categorization before the mutation.

The lower panel of Figure 5 shows the data for the maturation scenarios. Judging the young creature to be in the initial category (N = 112), and the mature creature to be in the final category (N = 94) are shown on the extreme left and extreme right of the figure respectively – and show the same phenomenalist pattern of justifications mostly referring to appearance and behavior. Those occasions on which the young creature was allocated to the final category (N = 16) showed a major use of the cause of change as justification. Clearly the fact that the process was one of natural development was of key importance to judging the immature form as being really in the final category. By contrast on occasions when the Origin essentialist position was taken of classifying the adult form in the initial category (N = 34), the cause of change was not often mentioned, but deep features (including offspring) were more prevalent.

Examples of justifications for different patterns of response were:

Phenomenalist post-mutation: "it looked and acted like a snake but had lizard offspring"

Origin essentialist post-mutation: "toxic contamination just changed what it looked like, not what it actually was"

Rips essentialist post-maturation: "since it's developmental, it must have always been a rabbit"

Origin essentialist post-maturation: "I still think it was a rhino – [it] had rhino offspring".

As in Experiment 1, the data were examined to see how many participants restricted themselves to justifications referring only to appearance or behavior. There were 9 of 16 such participants in the Typicality condition, 4 in the Absolute Membership condition and 5 in the Graded Membership condition. All showed the phenomenalist pattern of categorization. Therefore about half of the phenomenalist categorizers may have been truly phenomenalist, whereas the others were apparently also taking offspring or cause factors into account.

In sum, the justifications provided 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 does not alter an animal's category was justified through mention of offspring and the cause of the change. Judging that an immature form is really in the adult category was justified by appeal to the cause of change, while 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 found that a sizeable majority in the categorization conditions categorized on the basis of appearance and behavior alone. Collecting typicality and categorization judgments within-participants in Experiment 1 does therefore appear to have encouraged essentialist responding. Individual analysis confirmed that people tend to adopt one of four patterns of responding. Analysis of the justifications was consistent with the idea that for mutation scenarios sometimes people focus on the fact that appearance and behavior have changed, while at other times they focus on the fact that the offspring have *not* changed. For maturation scenarios, sometimes animals were classified on the basis of appearance and behavior, sometimes there was a focus on the analogy with natural dimorphism and both forms were classified in the adult category, while at other times focus on the offspring led to both forms being placed 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, they 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 for the Standard Condition of Experiment 3. First, 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—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. Another difference from the previous experiment was the inclusion of 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 2 was to investigate a factor commonly held to affect essentialist categorization, namely, the animal's offspring (e.g., Gelman & Wellman, 1991; Pothos & Hahn, 2000; Rosengren, Gelman, Kalish, & McCormick, 1991; Springer, 1996). Given that offspring are considered 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 the offspring was omitted. In the mutation scenarios of Experiments 1 and 2, an essentialist response could indicate one of two beliefs. One possibility is that a person believes that the germ-line of the creature is unchanged (because of the information about offspring), and, believing that genetics determine category kind, they therefore judge that the creature has not changed kind. Alternatively there may be people with a deeper essentialist belief – namely that a creature can never change its kind through an external cause such as toxic contamination. It is what it was born as, come what may. The Reduced Condition provides a test of these two alternatives. To the extent that essentialists are responding to genetic information inferred from offspring, in the Reduced Condition they should be more willing to entertain the possibility that the essence of the creature has also been changed by the mutation, and so respond with 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 had a 2 (Condition: standard, reduced) x 2 (Cause: mutation, maturation) x 2 (Phase: initial, final) mixed design, with Condition between-subjects, and Cause and Phase within. 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. Participants were sampled in the same way, but those in the Standard Condition were run prior to 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 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 their judgments. Finally, whereas Experiments 1 and 2 used paper booklets, the data for Experiment 3 were collected via 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 so 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 creatures before the change were judged highly similar to and typical of the initial category (*M* between 1.31 and 1.43, *SE* = 0.10), and after the change 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 simply followed the description of the creatures' appearance and behavior. The effect of Phase was highly significant for both measures in a 3-way ANOVA and no other effects or interactions approached significance.

Mean probability of 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, Reduced) x 2 (Cause: mutation, maturation) x 2 (Phase: initial, 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 $F'(1,105) = 4.56$, $p < .05$). These two effects also interacted ($F_P(1,91) = 5.1$, $p < .05$, $F_i(1,15) = 42.6$, $p < .001$). Before the change 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 where offspring information was omitted.

Confidence ratings showed a significant effect of Phase (Min $F'(1,106) = 7.8$, $p < .01$) and a significant Phase by Condition interaction (Min $F'(1,43) = 9.6$, $p < .005$). As seen in Figure 7, 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, there was greater confidence for the Reduced condition ($M = 4.4$) than for 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. 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 x Phase design. The criterion for classifying an individual now required that at least 6 of the 8 responses, in each of the four experimental cells must follow the expected pattern). Of the 46 participants in the Standard condition, 37 were Phenomenalists, 5 were Origin essentialists, and 1 was a Nominal essentialist. None exhibited the Rips essentialist pattern, and only one (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 could be expected from the analysis of the means, 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 one gave consistent phenomenalist responses, 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, 2-tailed).

Discussion

Contrary to our expectations, very few participants in the Standard condition endorsed essentialism. Prior to the 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. Categorization of the changed creature in the final category went up, 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 (rather than taking the view that an external cause can *never* change the kind of an animal).

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 were 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. Other differences were that the category judgments were within-scenario and dichotomous for Experiment 3. Both of these changes were intended only to clarify the categorization task. We can think of no reason why either one should reduce essentialist responses. A final pair of factors which could be considered is that the experiments were conducted on different student populations, and that while one experiment was done with paper and pencil the other was computerized. Differences in participant motivation or understanding of the task are therefore also possible accounts of 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. Taking the Standard condition together with the two earlier experiments, of the 110 participants who categorized scenarios, 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 notion of causal homeostasis. Pre-transformation there was no reason to doubt the animal's category, since it looked and acted one way and in fact its offspring looked and acted the same way. But after the transformation, the animal's appearance and behavior contrasted that of its offspring. As there was evidence both for and against categorization in the final category, confidence was lower. In the Reduced Condition, where the contrary information was removed, then confidence was restored.

General Discussion

The first experiment replicated Rips's (1989) seminal dissociation of typicality and categorization. While 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 (38%) of categorizers 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 now 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 Experiment 3 (Standard Condition) by asking about both the initial and 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 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 these 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 the 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 3 different patterns in the maturation condition. Of those participants across all studies 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 the phenomenalists agreed with the latter 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 between 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 in accordance with its appearance? In the introduction we identified two main hypotheses concerning people's beliefs about natural kinds. On the causal homeostasis view, a creature is the kind it is because of deeper 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. Given that post-transformation justifications often referred to deeper features in addition to appearance, but rarely referred to the cause of the change, there was evidence that changed categorization was based on the belief that deeper causal processes had also changed. The increased confidence when offspring information was removed (Experiment 3), also suggests that people were considering information other than appearance and behavior, but felt that the inferred deeper features outweighed the evidence of the offspring. Even though the offspring presented good evidence that the germ line was unaffected, the change in deeper structure was sufficient to change the kind to which the creature belonged. A parallel example would be the case of a person undergoing sex reassignment surgery. While the person's genetic essence is presumably unaffected, their friends and colleagues are asked to reclassify them to a different sex category. As with our results here, 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 consider it most appropriate to name a thing on the basis of the conceptual category to which it is most similar. 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 in its containing H_2O, but also to its origin and function. Naming something on the basis of its appearance certainly makes good sense for pragmatic reasons, for example if one wanted to point it out to someone else (Braisby, 2004). In folk taxonomy there are many cases of species that have the same name although they are not in the same biological category – for example the Jerusalem Oak and the Poison Oak are not in the oak genus *Quercus*. So Phenomenalist responders may have been reasoning pragmatically by naming the animals on the basis of similarity to other known animals. Understanding the categorization task as one of "finding the best name" for something would lead to changes in the category for both mutation and maturation. It is possible that if instructions were to explicitly contrast names for things against the underlying nature of things, then 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 naming categories. However the issue deserves further exploration.

In sum, the essentialist pattern of categorization by which a creature retains its kind through accidental or natural transformations, is by no means the only way in which people react to stories of biological metamorphosis. In the majority of cases people are more likely to see a creature that changes its appearance and behavior as having also changed its category. On the basis of the justifications given by participants we have argued that this change in category is based on a causal homeostasis view of how natural kind categories are determined. People infer from the change in appearance and behavior that the deep causal principles of the creature are now those of a new category. The response pattern taken on its own is however also consistent with people deciding to categorize creatures on the basis of their appearance and behavior, without any regard to internal causes or deeper knowledge. It remains an important issue 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.

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* (p. 87-107). New York, NY: 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: MIT Press.

Braisby, N. (2004). Similarity and categorization: Getting dissociations in perspective. *Proceedings of the Twenty-Sixth Annual Conference of 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 and 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). Categorization, causation and the limits of understanding. In H. E. Moss and J. A. Hampton (Eds.) *Conceptual Representation,* pp. 663-692. Hove: 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: 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.* 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 MS, New York University, May 2006.

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). Individual differences in essentialist beliefs. Poster presented at the 47th Annual Meeting of the Psychonomic Society, Houston, November.

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: 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 (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.

Author Notes

We are indebted to Lance Rips for graciously providing copies of his original materials. We also thank Sam Glucksberg, Yevgeniya Goldvarg, Uri Hasson, Phil Johnson-Laird, and Mary Newsome for helpful comments during the early stages of this research. Address correspondence concerning this article to Professor James Hampton, Department of Psychology, City University, London EC1V OHB, hampton@city.ac.uk

Footnotes

1. Gelman has a second form of essentialism – “sortal essentialism” – that more closely resembles our use 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” 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, while the offspring 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. The heading (e.g. Worm-snail) was not included. The element [*] was “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. In the Reduced Condition of Experiment 3, the final sentence was omitted.

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.

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.

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.

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.

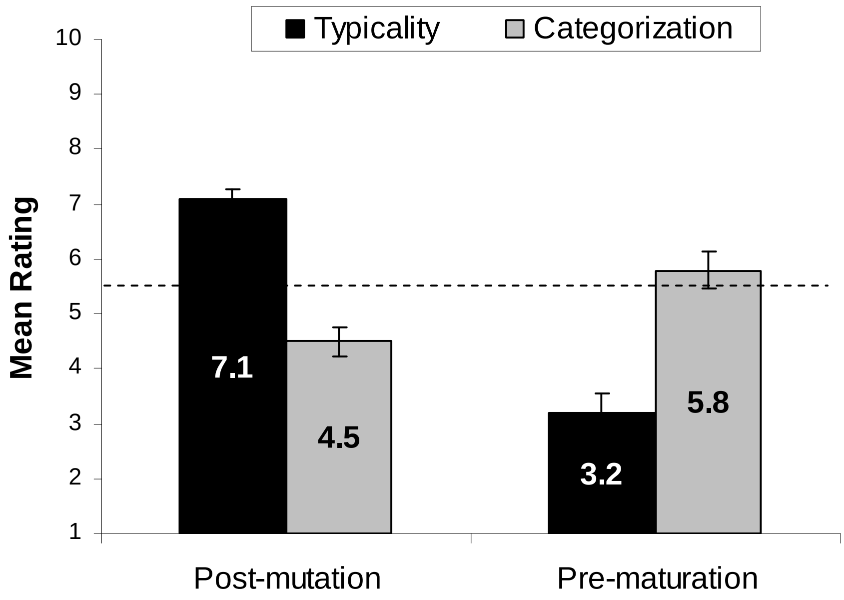
Table 1.

Proportions (and frequencies) of participants exhibiting each of the defined categorization patterns in each Experiment.

	Mutation		Maturation		Expt. 1	Expt. 2	Expt 2	Expt. 3	Expt. 3
Categorization Pattern	Pre-	Post-	Pre-	Post-		Absolute	Graded	Standard	Reduced
Phenomenalist	Initial	Final	Initial	Final	.38 (12)	.50 (8)	.69 (11)	.80 (37)	.98 (46)
Rips Essentialist	Initial	Initial	Final	Final	.16 (5)	.12 (2)	.00 (0)	.00 (0)	.00 (0)
Origin Essentialist	Initial	Initial	Initial	Initial	.28 (9)	.19 (3)	.12 (2)	.11 (5)	.02 (1)
Nominal Essentialist	Initial	Initial	Initial	Final	.09 (3)	.00 (0)	.06 (1)	.02 (1)	.00 (0)
Total proportion of consistent patterns					.91 (29)	.81 (13)	.87 (14)	.93 (43)	1.00 (47)

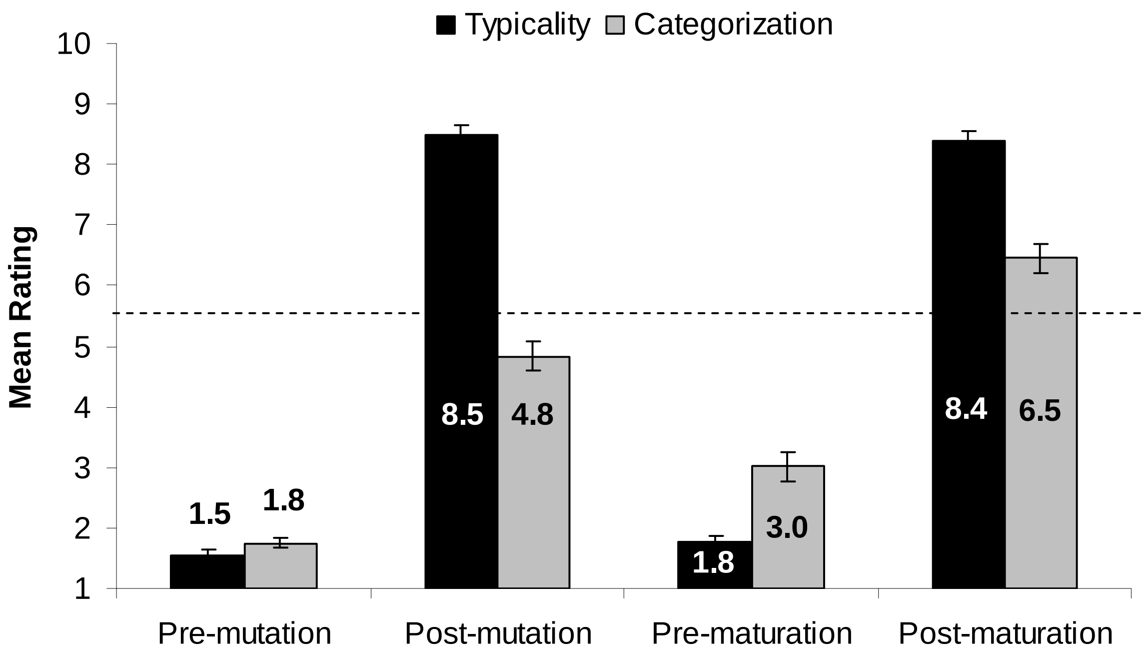
Note. “Initial” indicates a judgment that the animal is a member of the initially described category, and “Final” indicates a judgment that it is a member of the finally described category.

Figure 1. Typicality and categorization ratings (M ± SE), Rips (1989).



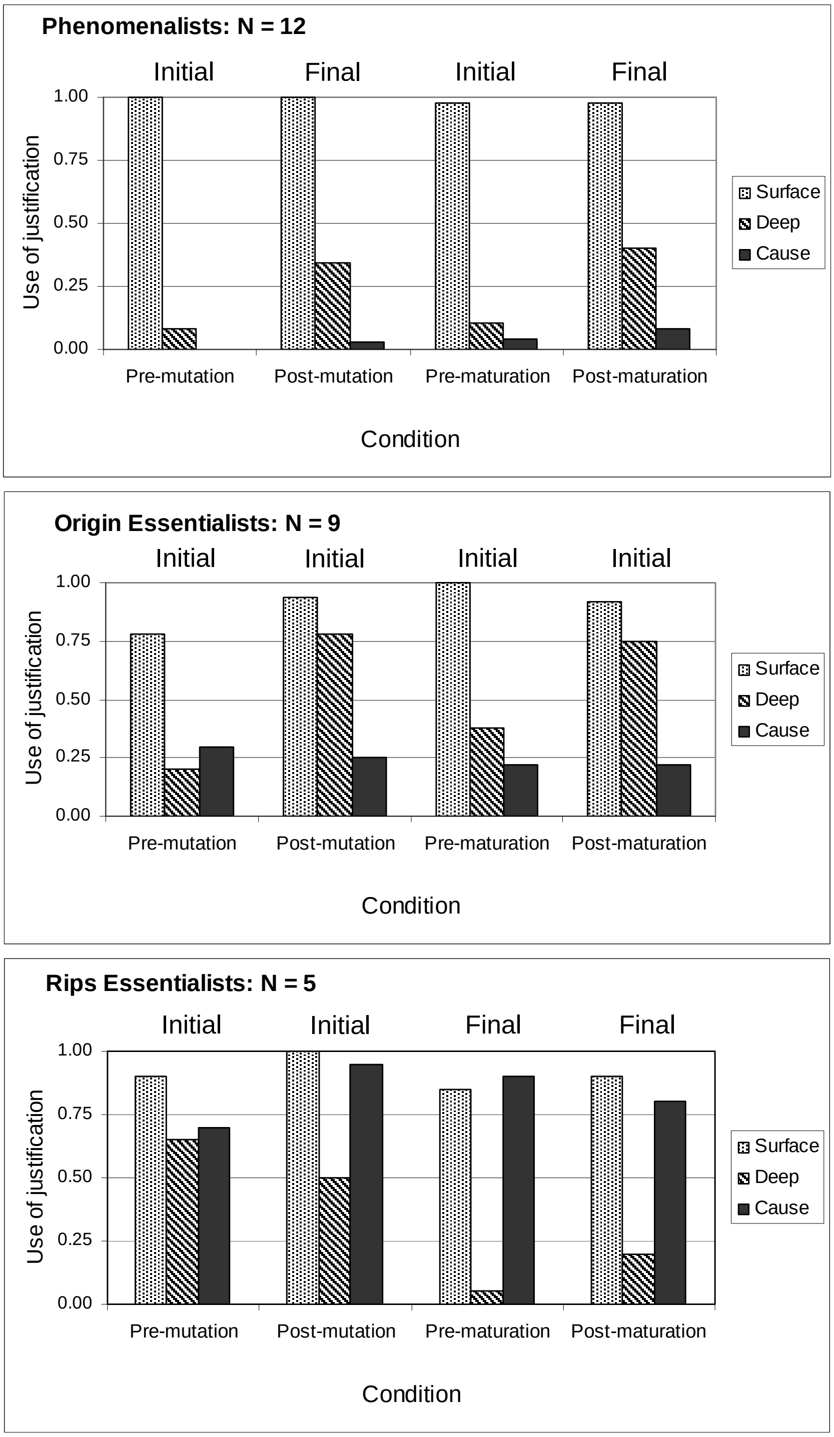
Note. Data are estimated from Figures 1.5 and 1.6 in "Similarity, typicality, and categorization," by L. J. Rips, 1989, pp. 21-59 in S. Vosniadou and A. Ortony (Eds.), *Similarity and Analogical Reasoning*. We have reversed Rips's original scale, so that 1 = initial category and 10 = final category. The dashed line shows the midpoint on the scale.

Figure 2. Typicality and categorization ratings (M ± SE), Experiment 1.



Note. Scale: 1 = initial category, 10 = final category. The dashed line shows the midpoint on the scale.

Figure 3. Use of justifications by participant groups in Experiment 1.



Note: The labels "Initial" and "Final" above each set of bars refers to the category to which the scenario was assigned.

Figure 4. Typicality and categorization ratings (M ± SE), Experiment 2.

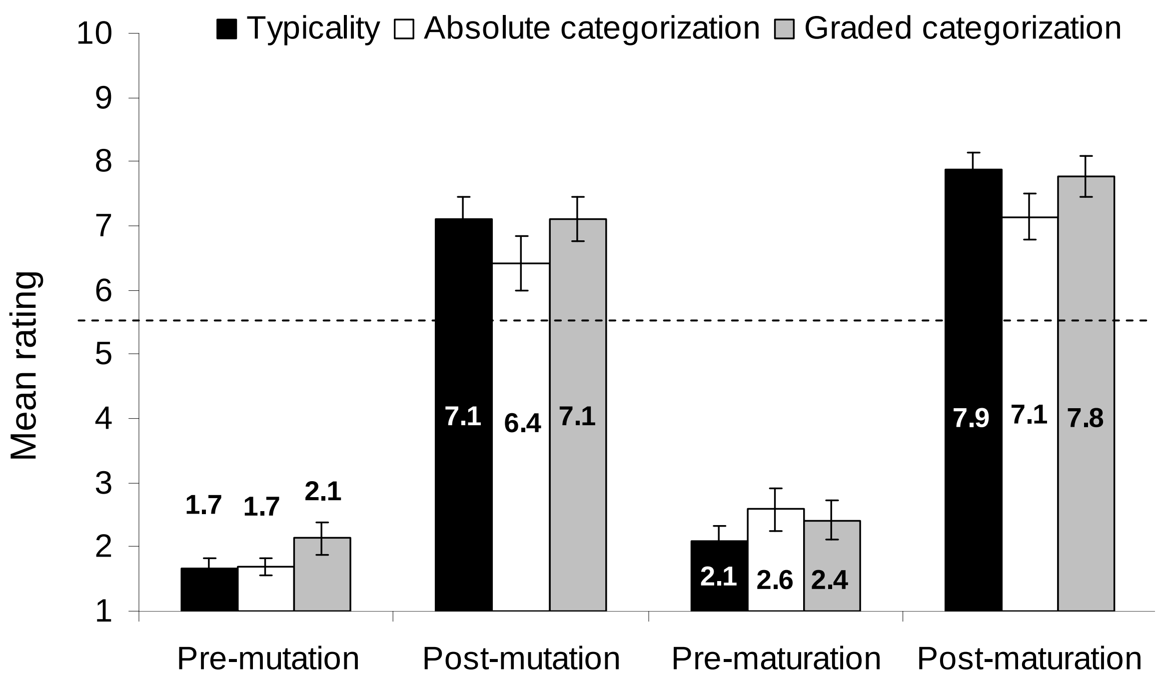
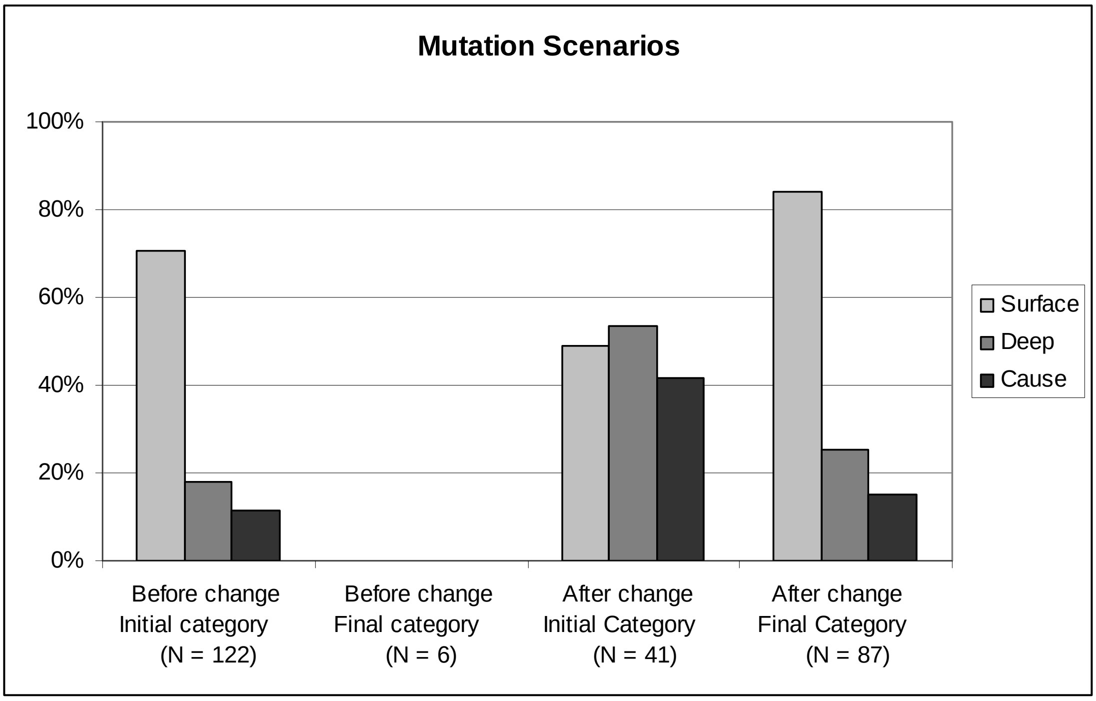
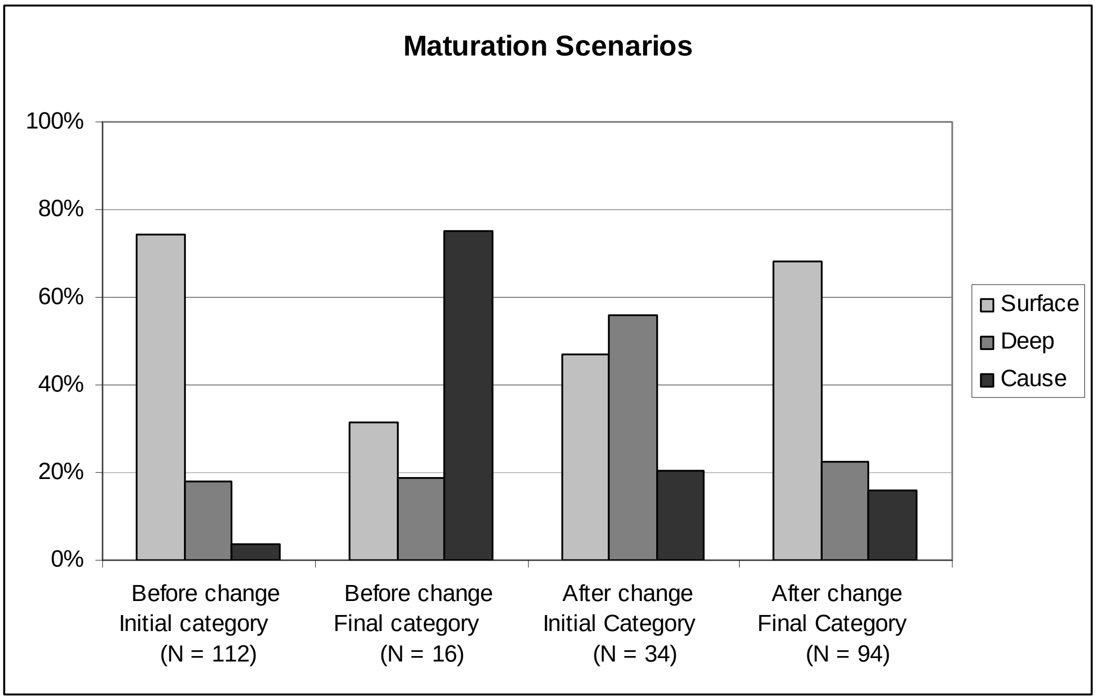


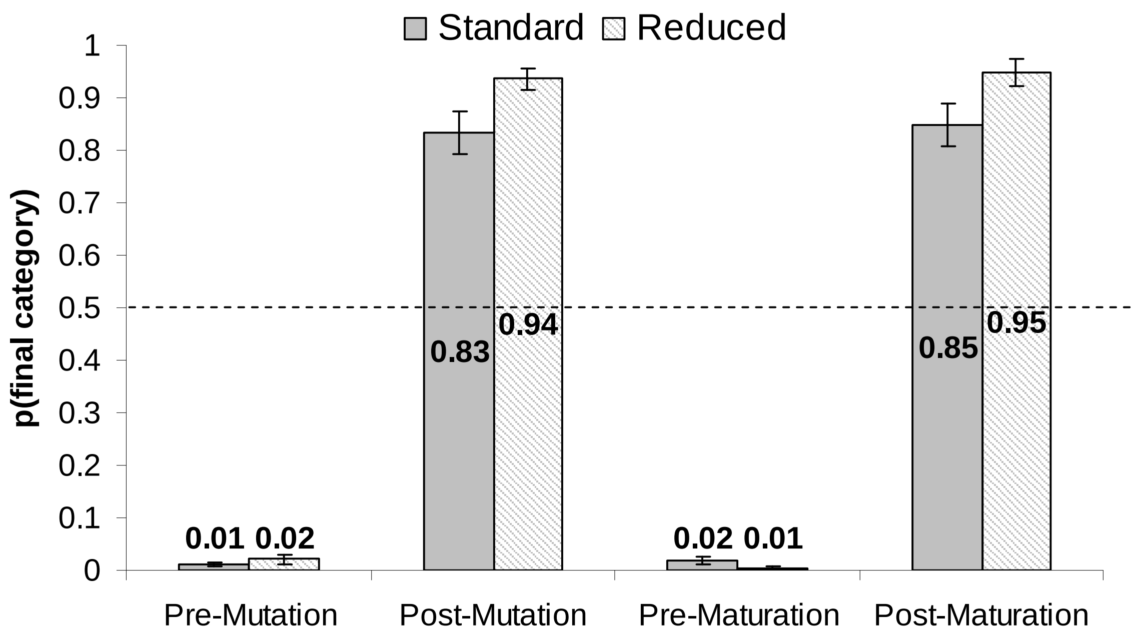
Figure 5. Percentage use of justifications in Experiment 2, broken down by category chosen and scenario type.





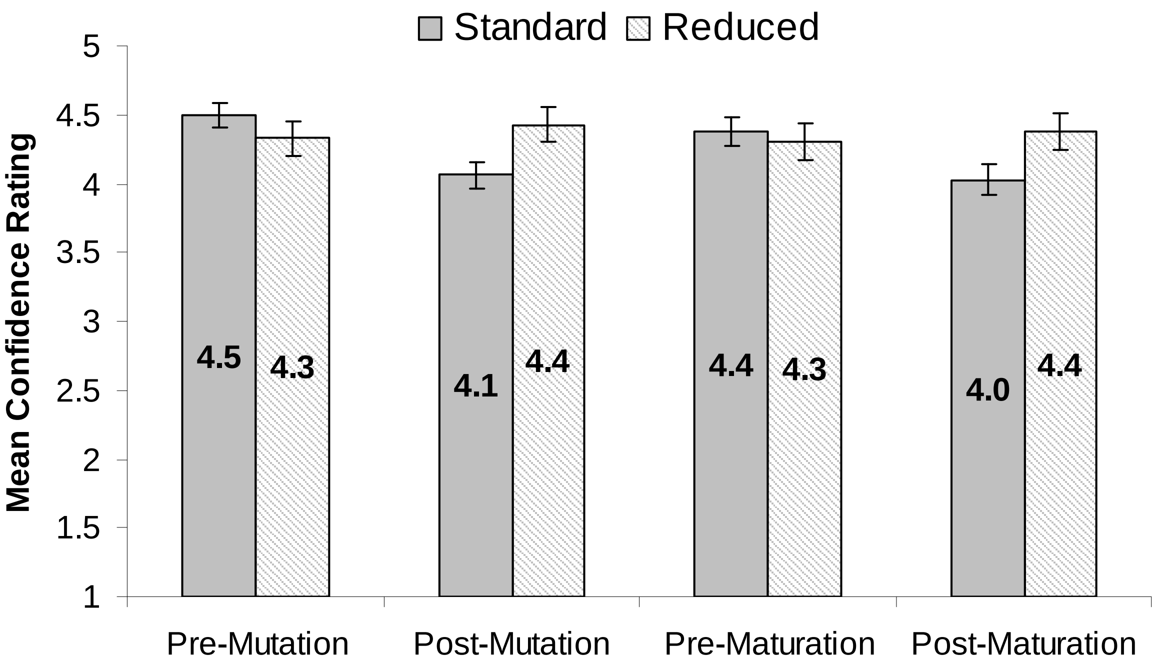
Note: there were insufficient data to plot for the choice of Final category before a mutation.

Figure 6. Categorization probabilities (M ± SE), Experiment 3.



Note. The Standard Condition included information that the animal's offspring resembled its initial category, whereas the Reduced Condition excluded that information. The dashed line shows the midpoint on the scale.

Figure 7. Confidence ratings (M ± SE), Experiment 3.



Note. The Standard Condition included information that the animal's offspring resembled its initial category, whereas the Reduced Condition excluded that information.