

# In Defense of Psychological Essentialism

Fei Xu & Mijke Rhemtulla  
fei@psych.ubc.ca, mijke@psych.ubc.ca  
Department of Psychology, 2136 West Mall  
University of British Columbia  
Vancouver, B.C., Canada V6T 1Z4

## Abstract

This paper attempts another defense of psychological essentialism (Strevens, 2000, 2001; Ahn, Kalish, Gelman, Medin, Luhman, Atran, Coley & Shafto, 2001). Using evidence from adults' and children's understanding of artifact concepts, we argue that the notion of essence does play a role in everyday reasoning and inference. Furthermore, there is also some corroborating evidence from the biological domain, contra Strevens.

**Keywords:** psychological essentialism, concepts, artifact concepts.

A recent exchange between Strevens (2000, 2001) and Ahn et al. (2001) has forced both psychologists and philosophers to take a closer look at the very idea of "psychological essentialism." Strevens (2000) laid out clearly the three versions of psychological essentialism as well as his alternative *minimal hypothesis*. All three versions of psychological essentialism posit essences at the center of children and adults' representational scheme for making inferences about observable properties of natural kinds. The minimal hypothesis suggests, contra psychological essentialism, that the notion of essence does not play a role in our everyday inferences. Instead, Strevens argued that *K-laws*, i.e., kind-based causal laws that guide our inferences about observable properties, are enough to back our inferences about kind membership. That is, to make inferences about *tiger* membership, one needs only know that "being a tiger" gives rise to tiger-properties such as stripes and ferocity, independent of any tiger-essence that may or may not exist.

We agree with many of the arguments made in Ahn et al. (2001) and we echo some of them below. This paper attempts another defense of psychological essentialism. Using evidence from the understanding of artifact concepts, we argue that 1) one source of evidence – data on the understanding of artifact

concepts by adults and children – may support the idea that essences do enter into our daily reasoning and judgment, 2) the fact that children may show understanding of essences of artifact kinds earlier than natural kinds strongly suggests that human beings have a natural inclination to look for essences, and 3) the evidence against the use of essences in reasoning about biological kinds may not be as strong as Strevens suggests. Inspired by the exchange between Strevens and Ahn et al., we also suggest that empirical investigations of the nature of kind concepts and psychological essentialism may take a different, more direct route.

The key claim in Strevens (2000) is that although representationally we may posit an 'essence box' ("essence of tigerhood") between category identity ("being a tiger") and observable properties ("ferocity" and "stripes"), as far as psychological processes like categorization and inferences are concerned, the notion of essence does not play a role at all. Inasmuch as our reasoning goes like this: "That thing is an X, therefore it must have X essence, therefore it shows X-properties," Strevens argues that we may as well get rid of the middle step, thus, "That thing is an X, therefore it has X-properties." Patterns of categorization and inferences can all be explained in terms of K-laws: Our understanding that kinds are causally connected with observable surface properties dictates our categorization and inference patterns. This is all the machinery we need. Let's forget about essences as playing any explanatory role in our everyday reasoning.

We do, however, need some way of getting to the first inference, "that thing is an X." Advocates of psychological essentialism would suggest that some observable properties (e.g., large, four-legged object capable of self-generated motion; roaring) lead to the inference that this thing is a tiger with tiger essence and other properties such as ferocity and stripes would follow. According to Strevens, the first set of observable properties would lead to the inference that this thing is a tiger, and given this kind membership

information, the other properties such as ferocity would follow. On one reading of Strevens, his views are perfectly consistent with psychological essentialism – he simply replaces the word “essence” with the word “kind.” In other words, “being a tiger” is different from “being striped” – the former identifies its kind (or essence, depending on one’s terminology) and the latter identifies an attribute that may or may not be important.

But let’s get back to the claim that the notion of essence does not play any role in everyday reasoning and categorization. The examples cited in this exchange were mostly from the biological domain. For instance, if something is a cow but was raised by a horse, is it going to be a cow or a horse when it grows up? Focusing on the biological domain is a sensible thing to do: the notion of essence may be most applicable to biological kinds (or more broadly, natural kinds) as some have argued (e.g., S. Gelman, 2003; Keil, 1989; Schwartz, 1979) and many studies have been conducted given this assumption. However, several psychologists have argued in recent years that artifact concepts also fall under the basic tenets of psychological essentialism (e.g., Bloom, 1998; Hall, 1995; Kelemen, 1999; Kelemen & Carey, in press). The suggestion is that the original intended function, or designer’s intent, is the essence of an artifact. According to psychological essentialism (e.g., Gelman, 2003; Keil, 1989; Medin & Ortony, 1989), when applied to a natural kind concept, the essence determines the surface properties we observe. This idea can be applied to an artifact kind as well: the original intended function determines the surface properties of an artifact, the actual uses it can fulfill, and its kind. For example, a chair is suitable for sitting because that is what its designer had intended. The fact that a chair has a flat surface and a back are properties that are appropriate for and determined by its intended function. A chair is a member of the kind chair because its designer had intended it to be a member of that particular artifact kind. A chair can also be used as a doorstopper or a ladder, but that is not why it came into being. It came into being because someone had intended to create a chair. Intuitively, we honor this distinction all the time, and would find it very confusing to refer to a chair differentially according to its current use (Imagine the following exchange. Person A: "The book is next to my alarm clock on the bedside table." Person B: "What bedside table?" Person A: "The one shaped like a chair." etc.).

If this analysis is correct, we can get back to the issue raised by Strevens (2000) and ask: Do adults and

children use original intended function – the essence of an artifact kind – when making judgment about kind membership? The answer seems to be “yes.” Rips (1989) showed that adults favored the original function of an artifact over its form in categorization. For example, if an object has the features of an umbrella but the original intent was to make a lampshade, adults judged that it was a lampshade. Matan and Carey (2001) found that an object made for containing tea but used for watering flowers was judged to be a teapot. Thus original intended function is weighted more heavily than current function or the object’s form. If we agree that original intended function is the essence of an artifact kind, these results provide strong evidence that adults do use information about essence in everyday reasoning.

In several developmental studies, Kelemen and colleagues have found that 4-year-old children are able to make kind membership judgments based on intended function. In Kelemen’s (1999) study, adults and children were shown pictures of novel artifacts that were designed for one purpose (e.g., squeezing lemons) but used by someone else for another purpose (e.g., picking up snails). When asked to judge what the objects were “for,” both adults and children favored the intended function. In another study (Kelemen, 2001), 4- and 5-year-old children were shown real novel objects and they also judged what the objects were ‘for’ based on intended function. Again, if we agree that intended function is the essence of an artifact kind, these data provide evidence that essences do play a role in everyday reasoning, even in young children who have not had any formal schooling.

Another source of evidence comes from Keil’s (1989) justly famous transformation studies. With artifacts such as coffee pots, adults and children were told that a coffee pot was modified to make a birdfeeder. Then they were asked to judge what the object was. The most common answer was “a birdfeeder.” Keil’s intention was to show that, unlike natural kinds, artifact kinds could change category membership easily because they do not have essences. However, if intended function really is the essence of an artifact kind, one might argue that the answer “a birdfeeder” was justified – after all, the adults and children were told that the coffee pot had been modified intentionally to make a birdfeeder! That is, someone intentionally used the original artifact as a base for designing some new artifact with a new function, and since intended function determines kind membership, the object has truly become a member of another kind via this change of intention (even when it

is not a change of intention on the part of the original designer).

Interestingly, the age (about 4 years) at which children first show an understanding of the essence of artifact kinds is younger than the age (about 6 years) at which they show sensitivity to insides or lineage, the best candidates for representations of essences in the biological domain. This is a bit peculiar since most agree that the notion of essence is more applicable in the biological and natural kind domain than in the domain of artifact kinds. One reason for believing in essences for natural kinds is because we believe that scientists may discover relevant facts that would inform us what the true essences are. For example, if scientists discovered that elephants were really robots controlled from Mars, their essence and our understanding of what their essence may be would change completely. Why is it then that children seem to have ‘discovered’ the essence of artifact kinds before they discover the essence of biological/natural kinds? We offer the following speculation. Children (as well as adults) have a natural inclination to find (and use) information about essences of kinds, contra Strevens (2000). In contemporary western society, discovering the essences of biological kinds is fairly difficult – since most of us do not live a life with lots of experiences with animals and plants, we have to rely on what others tell us. While parents may talk about birth and DNA fairly frequently, these are difficult concepts for children to grasp. For artifact kinds, on the other hand, we can imagine abundant evidence for intended function: children may well engage in games in which they have to make a novel artifact to serve a particular function, e.g., use playdough to make a bowl to hold water, or they have to bend a piece of wire to make a hook. Children are also very familiar with the frustrating situation of finding one's own creations misinterpreted: a child who has created a dinosaur out of building blocks will not take kindly to its being treated as a mountain. That is, children understand that their own intentions are supposed to determine the final object's kind identity and they are outraged by any misinterpretation of that intention. Such experiences may serve as the relevant ‘trigger’ for children who are looking for essences.

Even in the biological domain, we believe that the evidence is stronger than Strevens suggested (here we endorse an argument made in Ahn et al., 2001). One of Strevens' arguments was that internal essentialism does not work because children rely on both information about insides and information about lineage (parents and offspring) to reason about

biological kinds. Strevens was probably correct in saying that insides were not the only thing that mattered, thus the most literal reading of internal essentialism does not hold up. But let's consider the lineage information for a moment -- biologically, information about lineage is in fact a definitive clue to which essence (horse or cow) has been transmitted from generation to generation. Lineage is the mechanism by which essence is realized in each individual member of a kind. Thus we are inclined to suggest that children's reliance on lineage is actually evidence for psychological essentialism. In sum, we argue that an unusual source of evidence – understanding of artifact kinds – provides data against Strevens' minimal hypothesis. The fact that children and adults rely heavily on original intended function when reasoning about artifact kinds suggests that essence does play a role in everyday reasoning and kind judgment. Metaphysically, essences are least likely to exist in artifacts, as Strevens (2000) noted. Children and adults' tendency to look for essences in the least likely place, then, would seem to constitute a strong argument in favor of psychological essentialism. Furthermore, even in the biological domain, the use of lineage may be construed as evidence that adults and children rely on essence – in this case how the essences are passed on – when thinking about biological kinds.

We end on a methodological note. All parties agree that the basic tenet of psychological essentialism is the idea that essences are causally responsible for the surface features we observe. Many of the empirical studies to date have adopted the paradigm of pitting causally deep properties or information about essences against observable surface properties. As Medin and Ortony (1989) pointed out, discussions of psychological essentialism first appeared against the backdrop of theories of concepts that relied entirely on surface properties and similarity. The studies we have reviewed were often conducted with the goal of showing that adults and children reason according to causally deep properties and they are not simply bound by surface features and similarity. These studies, in our opinion, have shown convincingly that both adults and young children are not perception bound. Yet a more direct test of psychological essentialism would be to show that adults and children believe that essences are causally responsible for the surface properties observed. This, of course, is more easily said than done. But perhaps we psychologists should begin to devise such empirical tests (see also Fodor, 1998; Gelman, 2003). ]

## References

- Ahn, W., Kalish, C., Gelman, S.A., Medin, D.L., Luhman, C., Atran, S., Coley, J.D., & Shafto, P. (2001) Why essences are essential in the psychology of concepts. *Cognition*, 82, 59-69.
- Bloom, P. (1996) Intention, history, and artifact concepts. *Cognition*, 60, 1-29.
- Fodor, J. (1998) *Concepts*. Oxford: Oxford University Press.
- Gelman, S. A. (2003) *The essential child*. Oxford: Oxford University Press.
- Hall, D.G. (1995) Artifacts and origins. Unpublished manuscript. University of British Columbia.
- Keil, F. C. (1989) *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.
- Kelemen, D. (1999) The scope of teleological thinking in preschool children. *Cognition*, 70, 241-272.
- Kelemen, D. (2001) Intention in children's understanding of artifact function. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Minneapolis, MN.
- Kelemen, D. & Carey, S. (in press) The essence of artifacts: developing the design stance. In S. Laurence and E. Margolis (eds.), *Creations of the mind: Theories of artifacts and their representation*. Oxford University Press.
- Matan, A., & Carey, S. (2001) Developmental changes within the core of artifact concepts. *Cognition*, 78, 1-26.
- Medin, D. L., and Ortony, A. (1989) Psychological essentialism. In S. Vosniadou and A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 179-195). Cambridge: Cambridge University Press.
- Rips, L. J. (1989). Similarity, typicality and categorization. In S. Vosniadou and A. Ortony (Eds.) *Similarity and analogical reasoning* (pp. 21-59). Cambridge: Cambridge University Press.
- Schwartz, S. (1979) Natural kind terms. *Cognition*, 7, 301-315.
- Strevens, M. (2000) The essentialist aspect of naïve theories. *Cognition*, 74, 149-175.
- Strevens, M. (2001) Only causation matters: reply to Ahn et al. *Cognition*, 82, 71-76.