

Running Head: COMMENT ON RIPS, BLOK, AND NEWMAN (2006)

Sortal Concepts and Causal Continuity: Comment on Rips, Blok, and Newman (2006)

Mijke Rhemtulla and Fei Xu

University of British Columbia

(In press, Psychological Review)

Abstract

Rips, Blok, & Newman (2006) proposed that singular concepts, which support the tracing of individual objects across their existence, are governed by a principle of “causal continuity”. In three experiments, they purported to show that causal continuity does a better job than previously existing theories at explaining judgments of the persistence of individual objects. We make three points. First, the construct of “causal connectedness” entails a wide variety of different explanatory factors; calling them all “causal” has questionable explanatory value. Upon examination, there is little evidence that the ultimate basis for identity judgments in many cases is causal. Second, we suggest that causal knowledge is indeed important to identity, but that it is important in the context of sortal concepts: different causal information matters for different kinds of things. Finally, we consider the question of whether causal knowledge or sortal concepts are more fundamental to tracing individual identity; in other words, whether causal knowledge is strictly necessary for identity judgments. We appeal to research in developmental psychology that has begun to address this debate, supporting the primacy of sortal concepts. While there continues to be shortcomings of all theories of object persistence, it is not clear that the causal continuer theory brings new clarity to the puzzle.

Keywords: individual concepts, sortals, identity, causality, object persistence.

Sortal Concepts and Causal Continuity: Comment on Rips, Blok, and Newman (2006)

In recent years, the philosophical notion of sortals has come to light among psychologists and become the subject of some debate (Blok, Newman & Rips, 2005; Bonatti, Frot, Zangl, & Mehler, 2002; Hall, 1998; Liittschwager, 1995; Macnamara, 1986; Macnamara & Reyes, 1994; Rips, Blok & Newman, 2006; Xu, 1997, 1999, 2002, 2005; Xu & Carey, 1996; Xu, Carey, & Quint, 2004; Xu, Cote, & Baker, 2005). At issue is the role that sortal concepts, in particular, basic-level sortals such as PERSON, LION, and ICEBERG play in the psychological process of tracing individual identity. On the one hand, there are some compelling intuitive reasons to think they do. For example, consider how one might come to the decision that a bunch of cells attached to the inside of a woman should be considered a separate individual person, perhaps the object of considerable moral value? It would seem that human life begins at the moment that life becomes human; that is, the sortal concept PERSON (however it may differ among individuals) is the basis for a judgment of when personhood begins, and thus allows us to posit the existence of a new individual. To metaphysicians, a sortal “gathers up a class of things that survive certain sorts of change, come into being in a certain specific way, tend to be qualified in certain specific ways, and tend to cease to be in certain specific ways” (Wiggins, 2001; p.414). As a psychological theory, sortal theory allows people to have certain expectations about the persistence of individual objects in the world: caterpillars metamorphose into butterflies; young lions grow bigger until they reach maturity, and icebergs melt over time. As we learn facts relevant to specific kinds (whether by explicit teaching or by observation or inference), that information allows us to track the identity of objects over time (Hirsch, 1982; Macnamara, 1986).

In a series of evocative studies, Rips, Blok and Newman (2006) examined the role of causal reasoning in judgments of the persistence of individual objects. Their discussion tackled some of the central issues in the study of object identity, and proposed a theory of singular concepts (i.e., concepts of individual entities) that has causality at its core. Rips et al. argued that sortal theory is not adequate to explain judgments of individual persistence. Instead, they offered their *causal continuer* theory, which is based on the intuition that “the continuer of the original object must be a causal outgrowth of that original” (p. 7). Causal continuity captures the intuition that people think of causes as central to object persistence; like psychological essentialism, which posits a belief in deep causal essences, causal continuity suggests that what makes two entities identical with each other is not based on superficial similarity, but a deep causal connectedness.

This commentary consists of three main arguments. In Section 1, we hope to elucidate the inadequacies of the causal continuity approach: first, it is too general to provide any predictive or explanatory power. There are too many subtleties in the various ways people reason about persistence, that an appeal to general causal reasoning cannot account for them. Moreover, it may be the case that not all reasoning about individual persistence is causal. In Section 2, we will suggest that accumulated knowledge about particular sortals is essential to the process of reasoning about individual persistence. We will propose that rather than general causal information providing the basis for our reasoning about individuals of all kinds, causal information instead may be secondary to sortal concepts. Finally, in Section 3, we will look to existing research on cognitive development that addresses this debate.

Perhaps we should start by saying what we are *not* saying. First, we absolutely share the intuition that causal reasoning is of central importance to judgments of individual persistence.

Just as Rips et al. (2006) acknowledge that we cannot entirely get away from domain-specific knowledge, neither are we claiming that causal information is superfluous. Rather, the question, as we see it, is *how* causal knowledge is used. From the perspective of sortal theory, identity criteria differ for different kinds of things. Thus, different causal information is important for judging the identity of different kinds of things, and the same causal information may lead to different identity judgments depending on the kind of individual in question. For example, the knowledge that dogs grow and change through biological causal processes is part of our knowledge of the kind, DOG. If causal knowledge is subject to sortal concepts, then it is, in principle, impossible to take the sortals out of singular concepts. Rips et al. argue that, while sortals sometimes appear to guide our judgments about persistence, that is just because different kinds are subject to different causes. However, all kinds are subject to all types of causes; for example, people are affected by biological, intentional, chemical, and physical causes, yet only some of those are relevant to the task of tracing a person's identity. This is precisely where sortals may prove indispensable: representations of sortals are necessary to tell us which causes to pay attention to for each kind of individual. Second, we do not take issue with the second aspect of Rips et al.'s paper, that is, the decision-making model. This model is able to compute how participants will decide between two possible continuers, given their judgments regarding the identity of one potential continuer. Their model is admirably accurate. The success of that model, however, does not confer validity on their theoretical proposal.

Section 1: The Problems With Causal Continuity

What Is Causal About Identity Judgments?

Rips et al. (2006) describe three experiments that purport to give evidence for the role of causal factors in judgments of object persistence. However, it is not clear whether there is any

direct evidence that causal continuity underlies the observed judgments. In order to be able to make the claim that causal continuity is the factor that accounts for participants' judgments, Rips et al. would have had to provide, minimally, some measure of causal distance. Instead, they offer post-hoc descriptions, in each case, of how the observed results could have been due to causal reasoning. They rely on parsimony to make their argument: sortal theory fails to make the case for Experiment 1, and similarity fails to make it for Experiments 2 and 3. Therefore, they suggest, causal continuity is the better theory. While parsimony is undoubtedly a virtue of their theory, it is less certain that the umbrella framework that is proposed actually buys us anything in the way of predictive or explanatory power. To make this point, let us take a closer look at each case where "causal connectedness" is supposed to apply (e.g., Rips et al., 2006; Blok, Newman, & Rips, 2005; Liittschwager, 1995); we will see that this construct entails something different in each scenario where it is applied, resulting in a diverse array of criteria that are subsumed under the heading of causal connectedness.

Causal connectedness entails similarity of a proportion of particles. In Experiment 1 (Rips et al., 2006), participants read stories depicting Fred the lion, whose copied particles were combined in some proportion with particles from another lion or a tiger to create a new creature. Participants were asked to decide whether the resulting creature was or was not Fred. Results showed that participants' mean responses depended on the proportion of Fred-particles that made up the outcome creature. In line with this outcome, Rips et al. suggested that, "causal closeness in this experiment [depended] on the percentage of the copy's particles that [derived] from the original. In the stories, the copying machine is the causal mechanism that produces closeness by copying particles and transmitting them" (p.14). Thus, in this experiment, causal connectedness

looks like a compositional criterion, since it is measured by the percentage of the original's (copied) particles.

Causal connectedness entails spatiotemporal continuity of a discrete object. In Experiment 2 (Rips et al., 2006), participants were asked to rate the likelihood that a particular iceberg found at time 2 was identical to an original iceberg ("Sample 94") at time 1. Results showed that participants considered two assumptions in their decision-making: icebergs melt over time, and they tend to keep their shape when they melt. Based on these assumptions, they judged it to be unlikely that an iceberg at T_2 could be continuous with one at T_1 if it was bigger or very differently shaped than the original. Clearly, these judgments relied on causal knowledge; that is, knowledge about how ice behaves led participants to be able to guess whether a later iceberg was likely to be spatiotemporally continuous with the original iceberg. An iceberg that was larger at T_2 could not have been spatiotemporally continuous with the original iceberg (because ice does not tend to grow or expand, but melt), leading participants to judge that it was not the same iceberg. In this situation, causal knowledge acted as a cue to spatiotemporal continuity, which is the more fundamental criterion. If the two sources of knowledge were in conflict, we have little doubt that spatiotemporal continuity would win out (see Hall, 1998). For example, if participants had *watched* an iceberg grow in size, the spatiotemporal cues to identity would have overridden information about shape and size. As we alluded to in the introduction, general causal knowledge is not enough to account for identity judgments in the case of icebergs. Given a choice between two icebergs, one of which was likely spatiotemporally continuous and another which was not, participants could use causal principles to distinguish the two. But what if they were given a choice between the spatiotemporally continuous iceberg and the water that was once part of the original iceberg? Just as there is a

direct causal pathway between the original and the later iceberg, there is a direct causal pathway from the iceberg to the water that melted off it. Ultimately, we need more than general causal principles to account for the intuition. Here, then, we suggest that causal connectedness entailed a judgment of spatiotemporal continuity of the iceberg.

Causal connectedness entails continuity of proper name. In Experiment 3 (Rips et al. 2006), participants made judgments regarding the identity of streets and rivers following a fork in the road or river. A clear kind-related difference was observed: participants were more likely to infer the persistence of rivers than streets. Rips et al. argued that the relevant causal criterion was intentional; that is, participants reasoned that “authorities are more likely to redesignate streets at their convenience than to redesignate rivers” (p.23). The arbitrary re-designation of streets versus rivers seems to be a very distant cause, one that does not even impact the physical objects in question. Nevertheless, it is possible that this causal reasoning was behind participants’ identity judgments. Whatever their ultimate rationale, participants judged street and river identity based on whether or not the continuing forks were likely to have the same proper name as the original. In the case of people and other stimulus items across all these studies, proper names are a marker of underlying identity, rather than a determiner of that identity. If a person decides to rename himself, his pets, or other objects, we have the sense that just the name has changed, rather than the actual identity of the object. The identity of streets over space may be an exception: it may be that the identity of an individual street is completely determined by its proper name, such that arbitrary re-designation can sever identity. Thus, causal connectedness in this experiment entailed continuity of the original proper name.

Causal connectedness entails memory retention. Blok, Newman, and Rips (2005, Experiment 1) presented participants with scenarios wherein a man (say, Jim) involved in a

serious accident underwent radical surgery to transplant his brain into a new host. The host body was either a human body or a robot, and the story specified that either Jim's memories survived or did not survive. Participants tended to judge that Jim's persistence depended on the retention of his memories. As Blok et al. suggested, then, causal connectedness in this experiment entailed memory persistence.

Causal connectedness entails physical similarity. Blok et al. (2005) reported another series of experiments investigating individual identity judgments. In their Experiment 2, participants were shown pictures depicting scenarios where animals and artifacts were put into a particle transporter or copier. These machines worked by rearranging either the original or copied particles of whatever object was put into the machine, creating an output that was either identical to the original, belonged to a different but "related" kind (e.g., cat to dog), or belonged to a different domain (e.g., cat to boat). Blok et al. (2005) found that, when a related object was seen as the outcome, more participants judged it to retain its individual identity than when the outcome was unrelated. Unlike in Rips et al.'s Experiment 1, the percentage of particles never changed; all the particles from the original object were said to be rearranged (or copied then rearranged) into the outcome object. Despite this constancy, participants did not always judge the object to persist through the transformation. Instead, they judged identity to be dependent on physical similarity.

It is worth noting that Blok et al.'s (2005) justification of these results did not appeal to the cause of the transformation: "Bob's persistence should be more likely when the outcome object is related than unrelated, since it's easier to imagine Bob surviving as a dog than as a boat" (Blok et al., 2005; p.138). This effect was symmetrical; it did not matter whether the original object was an animal or an artifact. Apparently, it was also easier for participants to

imagine a boat surviving as a cup than as a cat. This symmetry suggests that, rather than considerations of “survival”, physical similarity was the real criterion for persistence judgments. It is also interesting to note that Blok et al. explicitly manipulated causal distance: half the participants saw depictions of a machine that rearranged the particles of the original object, and half saw a machine that first copied the original particles, then rearranged those copies. There was no effect of this manipulation. If causal reasoning is at the root of identity judgments, one might predict that adding an extra mechanism between the original object and the outcome should weaken their causal connectedness.

In another series of transformation experiments, Liittschwager (1995) revealed results that led to a similar conclusion as those of Blok et al. (2005). She showed children and adults transformations involving a magic fairy who transformed people across nine levels of change. These changes ranged from very superficial (e.g., clean boy to dirty boy) to more radical (baby to woman), to cross-basic-level kind (girl to cat), to cross-domain (woman to rain). Liittschwager’s results showed a gradual decrease in persistence judgments across levels of change: participants judged that when a person underwent a superficial change of role or appearance, her identity was retained. As the transformed object became more and more removed from humanity, participants’ attributions of persistence decreased steadily. In these experiments, the causal mechanism of change was a fairy, who effected transformations by tapping objects with her wand. Given that the causal mechanism was the same across all transformations (i.e., fairy magic), and the only difference from trial to trial was the physical form of the resulting object, it seems reasonable to conclude that participants relied on physical similarity to the original outcome in making their decisions. If causal reasoning was at the root of these judgments, then causal connectedness, here, entailed physical similarity.

Across all the experiments described above, then, causal connectedness entailed a staggering variety of things. It has encompassed the material composition of an animal, spatiotemporal continuity, the physical form of an animal or artifact, the retention of proper name, and the retention of memory in people. In each case, there is little or no evidence that the reasoning leading to these judgments was causally based. More worrisome, there is no sense in which the causal continuer theory predicted the observed results: certain causes come into play in some situations but not others, and there appears to be no principled way of making a priori judgments about which causes will affect the results in any given situation.

What Rips et al. have shown in their series of experiments is that everything that matters for individual persistence can be subsumed under “causal continuity.” This is hardly surprising, given the intuition that *causes*, ultimately, are responsible for everything. Why do helicopters hover? How do vacuum cleaners work? Though we may not know the answers at any level of detail, we do have the strong sense that some cause or other must be responsible. Likewise, we may have the sense that there are causes responsible for individual persistence. But the notion that causes affect individual continuity does not mean that people use causal information to make judgments about individual identity. Instead, it looks like people use information about physical appearance, spatiotemporal continuity, the retention of memories, etc. While it is true that we can tell causal stories about how objects get from point A to point B, Rips et al. have provided little evidence that these causal histories are the critical identity criteria, rather than various more obvious alternatives. In other words, subsuming the various factors under the umbrella term “causal continuity” does not provide us with any additional explanatory power if the theory cannot explain why, in each case, some particular identity criterion should be considered the “causal” one.

What Is Not Causal About Identity Judgments

We have attempted to show that, for causal continuity to successfully interpret the experimental data, it must be exceedingly non-specific; as such, we question its explanatory power. But there is another problem for causal continuity that may be even more damaging: It seems that there are some cases where the identity criteria that people use are simply not causal. We can think of two examples in particular. The first is Rips et al.'s (2006) Experiment 3, where participants judged the persistence of streets and rivers. It is possible, as Rips et al. argue, that participants took into account the "intentional causes" behind street and river naming practices. On the other hand, is it not more likely that they just used their knowledge of how streets versus rivers tend to be named? Streets, typically, do not make sharp turns; they intersect with other streets. Rivers, on the other hand, are far less likely to travel in straight lines. This knowledge is factual, not causal. This is not to say that people do not assume that these differences have a causal origin: it is very likely that we believe that there is a reason that streets and rivers are named differently, but that does not mean that we use causal reasoning in identity judgments as opposed to a simple piece of factual knowledge.

The second example comes from some of our own research. Rhemtulla and Hall (2006; Rhemtulla & Hall, 2005; Rhemtulla, 2005) found evidence that contextual factors can have a strong effect on judgments of individual persistence. Given a transformation-via-atom reassembler scenario very similar to the one used in Blok et al. (2005), our participants judged that individuals undergoing a transformation within a basic-level (e.g., Mallard duck to white duck) were significantly *less* likely to judge that the individual persisted (i.e., answer "yes" to the question, "Is this still [proper name]?") than if the transformation crossed the basic level (e.g., duck to cow).

This result challenges the conclusions of previous transformation studies reported above; in fact, if Blok et al.'s results are really evidence of causal continuity, then this result might be described as an anti-causal-distance effect. Participants in these studies were giving *lower* identity ratings to changes that resulted in a very similar-looking object than to changes that were more drastic. We interpret our data as resulting from a within-kind contrast effect: when faced with two different-looking exemplars of the same basic-level category, the differences between those exemplars become highly salient. In response, participants were unwilling to judge that both figures were instantiations of the same individual object. An example might make this surprising result intuitively clear: imagine a man, Sam, who is somehow transformed into a cat. Like many of our participants, you may have the intuition that the cat is still Sam, albeit in a different form. But imagine that, instead of being turned into a cat, Sam is turned into a man with a different appearance: perhaps you are more likely to believe that this new-looking man, whoever he is, is not Sam. We suspect this effect has more to do with the context of comparison than identity criteria, but if causal continuity finds support in similar transformation paradigms, then it must take seriously this result. We can think of no plausible reason that cross-basic transformations should be causally *closer* than those where basic-level kind membership is maintained.¹

SECTION 2: Sortal-Specific Causal Knowledge

In the previous section, we attempted to show that causal continuity, as specified by Rips et al. (2006), buys us no more than the idea that causality is a central notion in the way people reason about individuals. While we disagree with their theoretical formulation, we share the conviction that causal information plays an important role in judgments of individual identity. Unlike Rips et al., however, we think it is more likely that broad types of causal information

(e.g., biological vs. psychological) is subject to refinement in the context of particular sortal concepts. By this we mean that sortal concepts allow people to have specific ideas about which causes ought to matter to the identity of a person versus a dog, a chair, or a rosebush. Sortal concepts allow us to judge which causes to pay attention to in a given circumstance, depending entirely on the kind of individual in question.²

The Necessity of Sortal-Specific Knowledge

Let us consider the notion of ‘causality’ more carefully. Broadly speaking, we might describe three kinds of causes corresponding to three domains of core knowledge: naïve physics, naïve biology, and naïve psychology (Atran, 2000; Carey & Spelke, 1994; R. Gelman, 1990; Gelman & Hirshfeld, 1994; Medin & Atran, 2004). Each of these domains is associated with unique causal patterns, which are more or less relevant for different kinds of things (Coley, Hayes, Lawson, & Moloney, 2004; Coley, Medin, & Atran, 1997; Keil, 1994; Keil, 2003). Biological causality seems to be responsible for activities such as the growth and sustenance of plants and animals. Intentional causes act on people; they explain why people move around and do goal-directed things. Physical causes include chemical causes (e.g., the causes behind the melting of ice) as well as mechanical causes (i.e., Michottian contact causality). Not all causal knowledge is useful in the evaluation of individual identity. For example, Michottian causality is almost never going to bear on identity decisions, because this kind of causality is external to organisms and rarely affects object persistence. Individuals do not come into being, grow, change, or cease to exist due to being pushed around by external forces, unless that external force also causes some other internal change (such as an anvil falling on an animal resulting in its death).

Other kinds of causes matter for certain kinds of things but not others. In the domain of animals, internal biological causes are most important for identity judgments; knowledge of the results of biological causality helps us to predict how different plants and animals will grow and change over the course of their existence. When internal biological causes cease to play a role in the existence of a plant or animal (i.e., when it dies), we can say that that individual ceases to exist. What justifies this decision? When an animal dies, for example, it continues to be subject to many causes, such as the external chemical and biological causes responsible for the decay of its body. What knowledge do we draw on to decide that internal biological causes but not external causes should be relevant in determining the existence of animals and plants? Our point is that everything that happens to an animal is “causal”, yet only some of those causes matter for persistence. Another example may make this clearer. Take an artifact, for example an iron chair. Someone built this chair; as such we can say that the chair is the product of an intentional cause. If this chair is left outside for long enough, the entire thing may rust, rendering it completely different-looking. Despite this change to the substance of the chair, few people would judge that it is no longer the same chair. On the other hand, if the chair’s creator were to rearrange some of the pieces of the chair so that it functioned instead as a table, people would be more justified in saying that the table is a new individual entity. Both chemical causes and intentional causes acted on the chair, yet we have the intuition that only one of these causes matters for identity. The notion of “causality” must be informed by further domain-specific knowledge if it is to account for concepts of individual entities.

Sortal concepts provide the means for grouping individuals in the world into stable, coherent classes of objects. Consequently, we are able to acquire all sorts of real-world knowledge about things of a certain kind; some of this information may reflect our learning

about specific causal mechanisms, and some of it may reflect simple statistical regularities among members of a kind. Included in the knowledge we gain about kinds is specific identity criteria, which may be over and above that which can be gleaned from the simple accumulation of causal knowledge. What knowledge leads people to decide that an old man is the same individual as an infant 90 years prior, yet a year hence a corpse lying beneath that man's gravestone is no longer the same person? There is a causal connection between a man and his corpse, an ice cube and a puddle of water on the counter, and a mother and her child. Causal criteria alone cannot provide identity criteria to distinguish between two instances of one individual and two causally-connected yet non-identical individuals. In these cases, when a token ceases to fall under its sortal, its individual persistence is broken. Knowledge about particular sortals is necessary for a person to decide when a causal connection is enough to justify persistence, and when it is not.

Sortal knowledge can be acquired in many ways, including explicit teaching, observation, generalization, and induction. For example, a child presumably does not need to learn through separate processes that tigers, caterpillars and dodo birds grow by biological mechanisms: that knowledge can be generalized from one kind to others. The knowledge that icebergs melt, for example, does not have to come from explicitly watching icebergs melt, or being taught that icebergs melt. As an iceberg is really just a large chunk of the substance *ice* (at least according to our conception of it), and we know (from experience, or being taught) that ice melts, it follows that our concept ICEBERG inherits information from our understanding of ice.

What About the Transformation Studies?

Rips et al. (2006) agree that sortal-specific knowledge can play a role in identity judgments, perhaps contributing to the effects observed in their Experiments 2 and 3. Their case

against sortal theory rests on the patterns of identity judgments observed in their Experiment 1 as well as other transformation experiments in the literature (e.g., Blok et al., 2005; Liittschwager, 1995). In all these experiments, participants tended to judge persistence based on some measure of similarity which Rips et al. argue can be accounted for by causal distance. Contrary to the sortal theory, participants in each case did not use kind changes as a reliable cue to a change in identity. In order to make the claim that sortal concepts are fundamental to our reasoning processes, we must have some way to account for these transformation experiments. Here we have two points to make; the first is methodological and the second is theoretical.

Every study that found an effect of causal distance (i.e., Rips et al., 2006; Blok et al., 2005; Liittschwager, 1995) used a within-subjects design: each participant was asked to judge persistence at each level of transformation. In contrast, when a between-subjects methodology was used (i.e., Rhemtulla & Hall, 2005, 2006; Rhemtulla, 2005), the resulting pattern of identity judgments did not support the causal distance hypothesis; instead, the pattern that arose seemed to contradict the pattern that might have been predicted by causal continuity. As such, we suspect that the experimental design may have played a role in eliciting the observed pattern of judgments. In particular, within-subjects designs may encourage participants to adopt an implicit strategy of focussing on the variable that changes from trial to trial, and differentiating their responses accordingly. For example, in Rips et al.'s Experiment 1, where the percentage of Fred's particles varied from trial to trial, participants might have varied their judgments in response to that changing variable because it seemed like the most reasonable way to respond. We offer no independent evidence that this is the case, but it is perhaps one way of reconciling the different patterns of judgment observed across experiments.

Another reason to question the data from the transformation studies described above is that participants' judgments in those experiments do not seem to match our experiences with understanding the same sorts of scenarios presented in fiction. One motivation for some of this research has been the simple intuition that we allow individuals to change sortal in the context of fictional stories, such as fairy tales. In countless stories, the prince turns into a frog, or a beast, Cinderella's pumpkin turns into a coach, Gregor Samsa turns into a beetle, the Little Mermaid turns into sea foam, and we happily comprehend all of this. Interestingly, though, there does not seem to be anything like an effect of similarity (as a measure of causal distance) in our understanding of fairy tales: whether a person turns into a Beast, a frog, a beetle, or sea foam, we are equally credulous. We do not at all have the intuition that some of these stories are more believable than others, or that some preserve identity whereas others don't: the stories simply stipulate that identity is preserved, and we believe it, in the context of those fictions. This intuition leads us to suspect that identity judgments that do not correspond to real-world events are highly susceptible to the context of the fiction. Whereas in an experimental setting we see different judgments across transformational levels, if we were to test people's interpretations of familiar fairy tales, we might observe more uniform ratings of identity persistence, simply because the stories demand it for coherence. Of course, this is ultimately an empirical question. Our general point is this: whatever supports people's understanding of fictional and science-fictional transformations, it seems unlikely that experimental studies will reveal robust sophisticated reasoning mechanisms that can be analyzed. Given the malleability of participants' judgments based on the particulars of the experimental design, we may want to be wary of placing too much weight on these sorts of experiments.

Kinds vs. Individuals: A Clarification

What is the relationship between kinds and individuals, according to sortal theory?

Clearly, sortal theory predicts a strong connection between these, but it also clearly conceives of them as two very different kinds of concepts: sortals are kind concepts that come prior to individual concepts, which receive support from their corresponding sortals (Hirsch, 1982; Macnamara, 1986; Wiggins, 2001). In Rips et al.'s Experiment 1, particles copied from Fred were sometimes combined with other lion particles (where participants' ratings suggested that they tended to believe the outcome was a lion), otherwise they were combined with tiger particles (where most participants thought the resulting creature was no longer a lion). Rips et al. suggested that their finding that "factors affecting category membership [did] not necessarily affect decisions about individual persistence ... presents another puzzle for the view that identity conditions come from knowledge of sortal membership."(p.12). Here, we would like to clarify a potential point of confusion: Sortal theory *does* entail that every individual is fundamentally understood as a member of a particular kind, and thus predicts that an individual should not persist through a change in kind. On the other hand, sortal theory does *not* predict that an individual can undergo any process whatsoever and still retain its identity, so long as it remains a member of the same kind. For example, sortal theory does not predict that participants should be more likely to attribute individual persistence to Fred when the balance of particles come from another lion versus when they come from a tiger. In the most extreme example from Rips et al.'s Experiment 1, Fred was entirely replaced with particles from another lion; participants here clearly had no basis to judge that "still a lion" meant "still Fred"? There is not a one-to-one correspondence between kinds and individuals, and a sortal account of individual concepts does not simply equate individuals with kinds. In terms of Experiment 1, Fred and Calvin were two distinct individuals; sortal theory does not predict that they are interchangeable.

Recent discussion in the developmental literature (Bonatti et al., 2002; Xu et al., 2004) may have added somewhat to the confusion: in the is-it-one-or-two task, 12-month olds who saw a duck and a ball emerge in alternation from behind an occluder expected to see two objects revealed, and exhibited surprise at a one-object outcome. In contrast, 12-month olds who saw two different-looking balls (instead of a duck and a ball) appear on either side of an occluder were not surprised when only one object was revealed. In other words, it looks like infants thought that two objects of the same kind, no matter how different they looked, may have been the same individual. The main point of these studies was to show that when infants succeeded on the is-it-one-or-two task at 12 months, it was the sortal distinctions that supported their success. Surely infants eventually acquire sortal-specific criteria allowing them to infer, for example, two BALLS given a red and a green one but not necessarily two LEAVES given a red and a green one. Certainly, adults, who have had a lifetime of exposure to the way different kinds of individuals act in the world, have acquired these identity criteria. Young infants, though, who have not yet learned anything about balls or ducks, have no way of knowing that small balls do not spontaneously grow into large ones, or that red balls do not spontaneously change into green ones. To clarify, sortal theory does not suppose that the only knowledge we have of kinds is what makes instances of one kind different from others; we have, in addition, kind-specific knowledge about the characteristic modes of change that apply to individual members of different sortals.

Section 3: Developmental Evidence

Although we agree with Rips et al. that casual knowledge plays an important role, we believe that role is neither necessary nor sufficient to determine identity criteria. In Sections 1 and 2, we argued that causal information is not sufficient, because it is too unconstrained to give clear guidelines for when causal distance is too distant to allow persistence. In the following

section, we will argue that neither is causal knowledge necessary to make decisions about identity.

Some studies have shown that 9-month-old infants, who normally fail the “is-it-one-or-two?” object individuation task, succeed when given two contrastive labels. As the objects emerge from behind the occluder, one at a time, if the infant hears “Look, a toma!” and “Look, a fendle!”, she comes to expect two distinct objects when the occluder is removed on the test trials. Furthermore, providing the same label twice, two tones, two distinct sounds, or two emotional expressions does not help; thus, the effect may be specific to language (Xu, 2002). One interpretation of these results is that the distinct labels (in the form of count nouns) inform the infants that the objects belong to two sortals; therefore, there must be two object tokens in the event. Research with older infants and children has shown that children have a default assumption that the first word applied to an object will be a basic-level kind term, rather than a proper name, superordinate or subordinate kind label, or situational noun like “passenger” or “puppy” (Graham, Baker, & Poulin-Dubois, 1998; Hall, 1991, 1993; Hall, Waxman, & Hurwitz 1993; Soja, Carey, & Spelke, 1991). This research suggests that infants do not need to learn, for each new word, that it is a basic-level kind label; thus, an infant who sees an object paired with the word “fep” may immediately use that pairing to set up a kind representation. The infant needs only make the minimal assumption that the object on the stage is a *fep*, that a *fep* is a kind of thing, and that anything that is not a *fep* (e.g., something that is given the label “blicket”) cannot be the same individual as the thing that is a *fep*. Importantly for the current discussion, sortal terms do not convey any factual or causal information to the infant. The role of the sortal in this situation is simply a logical one: these are instances of two different kinds of things, therefore, there must be two individual objects. The infant has not learned any relevant facts

about ducks and balls, or about unfamiliar animals and artifacts. Instead, simply naming the two sortals is sufficient for the child to conclude that two objects are present. We take this as evidence that sortal concepts may come first, providing necessary support for infants' first singular concepts (e.g., Bonatti et al., 2002, suggested that infants have an innate concept, HUMAN, that supports singular person concepts such as MOTHER). Once these sortals are established, the child can accumulate causal and factual knowledge about them, enabling far more fine-tuned judgments of identity. It is clear, however, that there is much room for future research in this area. None of the existing research was conducted with the intent of distinguishing representations of sortal concepts from domain-general causal knowledge, an endeavour which could certainly inform the current debate.

Further developmental evidence may continue to shed light on this issue: we suspect that children learn many things about the characteristic activity of different kinds of things long before they understand the specific causes that cause those ways of being. For example, children might learn that plants grow, snowmen shrink, and chameleons change colour before they understand the relevant physical and biological causes that act upon those individuals. Indeed, it would be surprising if children could not trace the identity of individual entities long before they had learned about the mechanisms of biological and physical causality. Along this line of argument, Rosenblit & Keil (2002) have suggested that even adults drastically overestimate the extent of their own causal knowledge. While we no doubt appreciate that there are causal processes involved in the persistence of individual entities, even adults may not have the knowledge required to rely on understanding specific causal mechanisms in determining identity. Developmental research to date has suggested that even young infants have some understanding of causality; for example, recent research with seven and a half month old infants suggests that

they understand Michottian contact causality well enough to expect that an object will only cause another to move if it makes contact (Kotovsky & Baillargeon, 2000; Leslie & Keeble, 1987). It is difficult to see how an understanding of the causality of launching events would help infants to understand the much more complex internal causality that is responsible for the “causal continuity” of individual objects through space and time. Further research will be necessary to clarify the nature and depth of children’s understanding of causes that work to maintain individual identity.

Concluding Remarks

The finding that infants can acquire sortal concepts by hearing a label paired with an object may suggest that sortal concepts are dependent on language, rather than being dependent on the properties of the real world. Rips et al. (2006) pit this view of sortal concepts against their own view, on which people make use of worldly properties to judge identity – making causal continuity appear to be the more true-to-science approach. Sortal concepts, in contrast, may be seen as a wholly arbitrary way of carving up the world, regardless of its real properties. Contrary to this apparent dichotomy, most philosophers argue that sortal concepts are meant to reflect information about the real world, as it exists (e.g., Wiggins, 2001). The metaphysical sortal claim is that kinds exist: cats, icebergs, and rivers are natural kinds, and need only to be discovered. One consequence of this is that our concepts can change as we gain knowledge about the world. As Wiggins (2001) put it, “it is possible for what appear astonishingly like fs to be not fs, whereas the most improbable seeming specimens may turn out to be fs after all... It can also be a matter of prolonged and difficult inquiry gradually to improve currently accepted standards or conceptions of what it is to be an f.”(p.82). One could find out, for example, that frogs are, in fact, the same kind of animals as tadpoles – in such cases, the sortal account does not claim that

people should go on insisting that they represent two distinct kinds just because our prior concepts dictated that they were. On the contrary, Wiggins and others argue that sortal knowledge is based on scientific evidence, and that those concepts are thus being continually updated. In this sense, then, there is a deep similarity between the sortal account and the causal account proposed by Rips et al.: causal factors are undoubtedly at the root of every natural kind-specific change, and causal knowledge informs kind knowledge, just as kind knowledge helps us to know which causes are relevant.

Rips et al.'s (2006) paper is a thought-provoking exploration of the factors that guide our judgments of object persistence, and it highlights some of the most difficult issues that any theory of singular concepts must contend with. As Rips et al. note, discussions of causal reasoning and individual identity have been kept separate in the past. Their paper has made clear the need to bring these literatures together and examine the role of causal knowledge in reasoning about individual persistence. Rips et al. have proposed that a general understanding of causal forces naturally generates the relevant criteria for tracing individual objects in all domains, thus obviating the need for sortal-specific criteria. We have argued that in its current form, the causal continuer theory is too general to be useful in tracking the identities of various kinds of individuals. The theory must be made more specific such that it can generate predictions about the weight of relative kinds of causes in judgments of individual persistence. The stipulation that continuing identity simply entails causal continuity is insufficient to explain many identity judgments. The theory must specify how causal closeness is measured, how much causal closeness is necessary, and when and why some causal connections, though present, still lead to judgments of non-identity. It continues to be an open question whether, as the authors propose, individual identity can be and is traced using purely causal reasoning, in the absence of

sortal concepts. We hope that future research continues to open up this area of inquiry and shed more light on the structure of singular concepts, and their reliance on causal processes. We caution, though, that any theory of singular concepts must account for the pervasive domain-specific effects that are regularly observed in judgments of object persistence.

As discussed above, infants seem to acquire sortal concepts and with them the mechanism for individuation before they have any notion of causal processes. As such, we have argued that causal knowledge is not always necessary to support identity judgments. But this is not the end of the story by any means: as they progress through the world, children learn about the specific activity of particular kinds of entities: causal reasoning may well play a significant role in this learning. Importantly, though, causality alone is unable to do all the work: sortal concepts may set the background for further kind-specific learning, including causal learning, which in turn fleshes out our conceptual knowledge and allows for precise tracing of all different kinds of entities.

Footnotes

1. *Prima facie*, these data also seem problematic for the sortal view of individual persistence, because objects whose sortal membership was severed were more likely seen to retain identity than those who retained their sortal memberships. We interpret the experiment as an exploration of the role of comparison within and between kinds in singular concepts, perhaps based on the pragmatics of the situation, rather than a reflection of true identity criteria.

2. Note that we are not saying that tracing identity under a sortal is simply a matter of consulting a list of possible interactions to decide what will happen next. Such a view would imply that we have no capacity for causal reasoning at all. A commitment to sortal concepts as frameworks for organizing factual and causal knowledge does not commit us to storing every combination of possible interactions and scenarios that any kind of object could encounter. For example, what would happen if someone were to take a jackhammer to an iceberg? Of course people reason causally about such events: jackhammers break things into bits; icebergs are brittle, the iceberg is thus likely to break apart. But that causal reasoning does not lead directly to an identity judgment: which of the remaining iceberg-bits are identical with the original? Are any? Here, the sortal theory is committed to saying that the ultimate identity decision (i.e., where the iceberg is post-jackhammer, or if it exists) is dependent on the sortal ICEBERG.

References

- Atran, S. (1998). Folk biology and the anthropology of science. *Behavioral and Brain Sciences*, *21*, 547-609.
- Blok, S., Newman, G., & Rips, L. J. (2005). Individuals and their concepts. In W-k. Ahn, R. L. Goldstone, B.C. Love, A. B. Markman, & P. Wolff (Eds.), *Categorization inside and outside the lab*. Washington, DC: American Psychological Association.
- Bonatti, L., Frot, E., Zangl, R. and Mehler, J. (2002). The human first hypothesis: Identification of conspecifics and individuation of objects in the young infant. *Cognitive Psychology*, *44*, 388-426.
- Carey, S., & Spelke, E. (1994). Domain-specific knowledge and conceptual change. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 169-200). New York: Cambridge University Press.
- Coley, J.D., Medin, D., & Atran, S. (1997). Does rank have its privilege? Inductive inferences within folkbiological taxonomies. *Cognition*, *64*, 73-112.
- Coley, J. D., Hayes, B., Lawson, C., & Moloney, M. (2004). Knowledge, expectations, and inductive reasoning within conceptual hierarchies. *Cognition*, *90*, 217-253.
- Gelman, R. (1990). First principles organize attention to and learning about relevant data: Number and the animate-inanimate distinction as examples. *Cognitive Science*, *14*, 79-106.
- Gelman, S.A. & Hirshfeld, L. (1994). *Mapping the Mind: Domain Specificity in Cognition and Culture*. New York: Cambridge University Press.

- Hall, D. G. (1991). Acquiring proper names for familiar and unfamiliar animate objects: Two-year-olds' word-learning biases. *Child Development, 62*, 1142-1154.
- Hall, D. G. (1993). Basic-level individuals. *Cognition, 48*, 199-221.
- Hall, D. G. (1998). Continuity and the persistence of objects: When the whole is greater than the sum of the parts. *Cognitive Psychology, 37*, 28-59.
- Hall, D. G., Waxman, S. & Hurwitz, W. (1993). How two-and four-year-old children interpret adjectives and count nouns. *Child Development, 64*, 1651-1664.
- Hirsch, E. (1982). *The Concept of Identity*. New York, NY: Oxford University Press.
- Keil, F.C. (1994). The birth and nurturance of concepts by domains: The origins of concepts of living things. In L. A. Hirschfeld and S. A. Gelman (Eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture*. New York: Cambridge University Press.
- Keil, F. C. (2003). Categorization, causation, and the limits of understanding. *Language and Cognitive Processes, 18*, 663-692.
- Kotovskiy, L., & Baillargeon, R. (2000). Reasoning about collisions involving objects in 7.5-month-old infants. *Developmental Science, 3*, 344-359.
- Leslie, A., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition, 25*, 265-288.
- Liittschwager, J. C. (1995). Children's reasoning about identity across transformations. (Doctoral dissertation, Stanford University). *Dissertation Abstracts International, 55*, 4623B.
- Macnamara, J. (1986). *A Border Dispute: The place of logic in psychology*. Cambridge, MA: MIT Press.

- Macnamara, J. & Reyes, G.E. (1994) *The logical foundations of cognition*. Oxford University Press.
- Medin, D. L. & Atran, S. (2004). The native mind: Biological categorization and reasoning in development and across cultures. *Psychological Review*, *111*, 960-983.
- Rhemtulla, M., (2005). Concepts of individuals do not allow identity persistence within the basic level. Unpublished Master's Thesis.
- Rhemtulla, M., & Hall, D. G. (2005, June). Concepts of individuals do not allow identity maintenance within the basic level. Poster presented at the 2005 meeting of the Society for Philosophy and Psychology, Wake Forest, NC.
- Rhemtulla, M. & Hall, D. G. (2006). Basic-level kinds and individuals. Manuscript under review.
- Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: an illusion of explanatory depth. *Cognitive Science*, *26*, 521-562.
- Rips, L. J., Blok, S., & Newman, G. (2006). Tracing the Identity of Objects. *Psychological Review*, *113*, 1-30.
- Soja, N., Carey, S., & Spelke, E. (1991). Ontological categories guide young children's inductions of word meaning. *Cognition*, *38*, 179-211.
- Wiggins, D. (2001). *Sameness and Substance Renewed*. New York: Cambridge University Press.
- Xu, F. (1997). From Lot's wife to a pillar of salt: Evidence that physical object is a sortal concept. *Mind & Language*, *12*, 365-392.
- Xu, F. (1999). Object individuation and object identity in infancy: the role of spatiotemporal information, object property information, and language. *Acta Psychologica*, *102*, 113-136.

- Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition*, 85, 223-250.
- Xu, F. (2005). Categories, kinds, and object individuation in infancy. In L. Gershkoff-Stowe and D. Rakison (eds.), *Building object categories in developmental time* (pp. 63-89). Papers from the 32nd Carnegie Symposium on Cognition. New Jersey: Lawrence Erlbaum.
- Xu, F. & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, 30, 111-153.
- Xu, F., Carey, S., & Quint, N. (2004). The emergence of kind-based object individuation in infancy. *Cognitive Psychology*, 49, 155-190.
- Xu, F., Cote, M., & Baker, A. (2005) Labeling guides object individuation in 12-month-old infants. *Psychological Science*, 16, 372-377.

Author Note

Mijke Rhemtulla and Fei Xu, Department of Psychology, University of British Columbia.

We thank Geoff Hall for many hours of helpful discussion, and Henny Yeung and Amy Perfors for very helpful comments on an earlier draft of this manuscript. We also thank Lance Rips, Michael Strevens, and two anonymous reviewers for constructive criticisms and comments on an earlier version of the manuscript. This work was supported by NSERC and SSHRC grants to F. Xu.

Correspondence concerning this article should be addressed to Mijke Rhemtulla or Fei Xu, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, British Columbia, V6T 1Z4. E-mail: mijke@psych.ubc.ca or fei@psych.ubc.ca.