Human Development (in press)

The principle of persistence, Leibniz's law, and the computational task of object re-identification

Chris Fields

815 East Palace Ave. #14, Santa Fe, NM 87501 USA
fieldsres@gmail.com

Abstract:

Infant abilities to re-identify objects through time have been characterized in terms of a "principle of persistence" (R. Baillargeon (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science 3*, 2-13). It is argued that the proposed principle of persistence has neither theoretical content nor explanatory power, and that object persistence is itself not an explanatory principle but rather an effect to be explained. The computational task of object re-identification across gaps in observation is characterized, and shown to be computationally intractable and hence solvable only by means of heuristics. It is argued that understanding infant abilities to re-identify objects requires focused investigation of the object re-identification heuristics that infants actually deploy and their neurocognitive implementation.

Keywords: Categorization, Frame problem, Infancy, Object files, Trajectory recognition, Visual object segregation

Introduction

Over half a century of experiments and theoretical work have challenged Piaget's (1954) notion that infants begin life with negligible knowledge about the world with an alternative view of infants as displaying significant knowledge of the world from birth. A primary tenet of this alternative view is that even the youngest infants have a representation, possibly an innate representation, of what constitutes an *object*. Spelke (1994), for example, proposed that "young infants appear to make inferences about the hidden motions of inanimate, material objects in accord with three principles:

cohesion (objects move as connected, bounded units), continuity (objects move on connected, unobstructed paths) and contact (objects affect one another's motion if and only if they touch)" (p. 433; see also Spelke, 1998). "Hidden" motions are emphasized because experimental designs involving temporary occlusion provide the primary empirical basis for this claim; that infants also interpret the motions of non-hidden material objects as obeying these principles is to be understood. On the basis of an additional decade-and-a-half of experimental results, Baillargeon (2008) generalized Spelke's principles of cohesion and continuity to a principle of *persistence*: "The persistence principle states that objects not only exist continuously and remain cohesive, they also retain their individual properties" (p. 3). The "properties" that remain persistent depend on the properties that infants can attend to and represent as significant, which change as development proceeds; however, any properties that are salient to infants are treated, according to this principle, as persistent properties (see also Baillargeon, Li, Gertner & Wu, 2011; Baillargeon, Stavans, Wu, Gertner, Setoh, Kittredge & Bernard, 2012). The persistence principle implicitly defines an *object* as an entity that is persistent through time; Baillargeon's characterization of the persistence principle as an "innate idea," like Spelke's characterization of the principles of cohesion, continuity and contact as "initial knowledge," is effectively the claim that infants are born with a fundamental, presumably un- or pre-conscious conception of the world as composed of cohesive, bounded, time-persistent objects.

The idea that infants have an innate conception of objects as cohesive, bounded and persistent is not merely a philosophical claim: it has clear consequences for how infant behavior is viewed and interpreted. Many experimental designs commonly employed with infants would be difficult to comprehend without the assumptions that infants not only perceive and behave within a world of cohesive, bounded and persistent objects, but also that what are perceived to be objects by adult experimenters are also perceived to be objects by infants. Lou (2011), for example, recently tested whether 3-month-old infants would interpret the behavior of an apparently self-propelled box as intentional by measuring how long infants looked at scenes in which the box approached either a cone or a cylinder. Prior to testing, the infants were shown training scenes in which the box repeatedly approached the cone. Between training and test scenes, a curtain was lowered that hid the entire display. As the point of the training scenes was to enable the infants to infer that the repeatedly-approached cone was a goal of the box's motion, the design clearly requires the assumptions that the infants would perceive the box, the cone and the cylinder *as objects* during both the training and test scenes, and that they would treat each of these objects as *remaining the same object*, i.e. persisting, between the training and test scenes. If either of these assumptions is dropped and no more complicated assumptions are introduced (e.g. that 3-month-old infants generalize shape-dependent goals across multiple distinct but visually-similar boxes), Luo's interpretation of the looking-time results as indicating that infants were surprised when the box appeared to change its goal is logically unavailable; if the infants did not regard the box, for example, as an object, or if they regarded the box in the test scenes as a different object than the one(s) they saw in the training scenes, there would be no reason to for them to be surprised by the different motions executed by the box(es) in the different scenes.

The principles enunciated by Spelke (1994) and Baillargeon (2008) are, however, also considerably more than methodological prescriptions regarding the assumptions to be employed in the interpretation of experiments: both are explicitly claims about knowledge – expressed perhaps by capabilities – that infants actually possess. It is in this sense that the principles of cohesion, continuity and contact, and their extension to the principle of persistence, are *theoretical* claims. As such, they are intended not just to characterize but to *explain* the kinds of behaviors that infants exhibit, whether in an experimental setting or in ordinary life. This goal of explanation is evident throughout Spelke (1994) and Baillargeon (2008). Spelke concludes her paper with:

> A picture of cognitive development emerges from studies of initial knowledge. Humans are endowed with a number of systems of knowledge, each consisting of a limited set of principles capturing highly reliable constraints on a significant class of entities. Over the course of human development, each system of knowledge grows as the principles at its core are enriched by further, generally less reliable notions. In addition, distinct systems of knowledge come to guide an increasingly wide range of actions and come to be related to one another. Although studies of early development have not revealed the processes that enable children and adults to link distinct systems of knowledge to one another and to systems guiding action, they suggest situations within which psychologists may begin to study these processes. Studies of these processes, in turn, may shed light on an aspect of cognition that is perhaps unique to humans: our ability to extend our systems of knowledge into territory that lies beyond their initial bounds.
>
> Spelke (1994) p. 443

Baillargeon similarly regards the persistence principle as central to the explanation of infant behavior:

> Our account assumes that infants' representations of events – or physical representations – are initially impoverished but become richer with experience as infants gradually learn what information to include in order to better predict outcomes. Any information infants include in their physical representations becomes subject to a principle of persistence, which incorporates and extends the principles of continuity and cohesion…According to this principle, no object can undergo a spontaneous or uncaused change in the course of an event, be it winking out of existence; breaking apart; or changing size, shape, pattern, or color…According to our account, infants succeed in detecting continuity and change violations when they have included the necessary information to do so in their physical representations…Once included in the physical representation, the information becomes subject to the persistence principle.
>
> Baillargeon (2008) p. 3

Given that these principles are intended as explanations, one can fairly ask *how* they explain what they are purported to explain. Spelke states, somewhat paradoxically, that "a theory's explanatory value does *not* depend on the content it assigns to the initial state" (Spelke, 1998, p. 194; emphasis in original), claiming that what does count for explanation, aside from internal consistency, is "how well (the theory) accounts for the phenomena of development" including "the subsequent growth of knowledge" (both p. 194). What counts, in other words, is how the postulated initial knowledge or capabilities contribute functionally, for the infant, to subsequent learning and developmental processes, and hence inferentially, for the outside observer, to the theoretical explanation of subsequent learning and development. How, for example, does information acquired from the environment and included in a specific physical representation – say a visual representation of a box – "become subject to the persistence principle"? What does "becoming subject to the persistence principle" mean in terms of processes in an infant's nervous system? This question is a special case of the more general question that must be asked of any knowledge claimed to be innate, or indeed of any knowledge regardless of its provenance: how is the knowledge – in this case, the persistence principle – functionally implemented? It is this question of *implementation* with which the present paper is concerned. Without an answer to this question, it is entirely unclear what it means to invoke the persistence principle when saying, for example, that an infant *knows* that it is the *same* box approaching a cone.

While the persistence principle itself says nothing about its implementation, its application in practice reveals an implicit implementation-level constraint. The looking-time paradigm developed by Spelke (1985), and now employed routinely in studies of infant cognition, is specifically designed to only require that infants attentively *look at* a scene. Motor actions are not required by the paradigm, and are not desired during experiments. If the results of experiments that employ this paradigm – such as that of Luo (2011) – are to be explicable in terms of the persistence principle, then the principle must apply to the visual processing of information in an at least episodically action-independent way. Many of the experiments appealed to by both Spelke (1994) and Baillargeon (2008) employ the looking-time paradigm; hence it is fair to conclude that the principles of cohesion, continuity and contact, and the persistence principle are intended to characterize visual processing independent of motor actions. These principles must, therefore, be implemented by the infant visual processing system. They may also be implemented by the somatosensory or other perceptual systems, or by multi-modal perceptual binding, but these alternative and perhaps parallel implementations must not be required, at least episodically, by the visual-system implementation. Both claims of innate knowledge and experimental designs such as the looking-time paradigm that probe infant knowledge or capabilities in very restricted settings have been criticized from a variety of perspectives, not least of which is that key terms such as "knowledge" or "representation" remain ill-defined (e.g. Haith, 1998 or from a less-philosophical perspective Gredebäck & von Hofsten, 2007). Asking how the persistence principle is implemented not in general but rather in the context of a specific, anatomically well-defined and functionally well-characterized perceptual system allows this question of the meanings of terms such as "representation" to be directly addressed.

The remainder of this paper argues that while the principle of persistence appears to have both theoretical content and explanatory power, it in fact has neither. The next section "The implementation of object persistence" addresses the issue of theoretical content. It first reviews the functional architecture of the infant visual-processing system and its connections to the motor system, using data from both infants and adults. The often-unstated assumption that a representation of objects as persistent requires conceptual reasoning about object persistence is then criticized as introducing a straw man: there is no need to suppose that infants must understand metaphysical concepts to recognize objects. It is shown that the notion that infants or even adults employ "persistent object" as a fundamental concept or category from which more specialized concepts or categories representing particular kinds of objects are derived by the addition of particular features, although commonplace in ontological thinking, is theoretically vacuous for developmental psychology. From an architectural perspective, object persistence is an *outcome* of visual processing, not an explanation of how visual processing works. The notion that object features are persistent, a notion that embodies Leibniz's "Law" of the identity of indiscernibles, must be so modified by caveats and exceptions in order to remain useful in practice that it loses any effective force. The third section, "The computational task of object re-identification" further addresses the issue of explanatory power. It suggests that the principle of persistence might have been labeled more transparently as "the persistence effect" – i.e. as clearly being an observed regularity in need of explanation, as opposed to an explanatory claim. Explaining the persistence effect requires abandoning the comfortable idea that infants come into the world knowing about cohesive, bounded objects – and in particular, abandoning the idea that infants come into the world knowing what *count as* objects to adults – and confronting the difficult questions of how infants or even adults are able to segregate objects from their surrounding environments and able to determine when an object seen now is the same thing as an object seen a few seconds, minutes, days or years ago. Neither the object segregation problem nor the re-identification problem is computationally tractable; hence humans can "solve" either problem only heuristically. The paper concludes that understanding the development of object perception in infancy requires both understanding the heuristics that infants actually deploy to solve the object segregation and re-identification problems, and understanding how these heuristics are implemented by the developing neurocognitive system.

The implementation of object persistence

The architectural arena: The infant visuo-motor system

Before addressing the question of how the persistence principle is implemented by the infant visual system, it is worth reviewing the functional architecture of that system in adults and, as the two systems share representations, its coupling to the motor system. Functional studies in macaques indicated by the early 1980s that processing of visual stimuli is segregated into two anatomically distinct but mutually-modulatory "streams," a dorsal stream through the medial-temporal area (MT) that processes location and motion information, and a ventral stream through visual area 4 (V4) that processes shape, color and other static feature information; lesion studies in human adults confirmed this division (reviewed by Van Essen & Maunsell, 1983; Maunsell & Newsome, 1987). This functional-anatomical distinction was quickly linked to behaviorally-evident differences in the processing of motion ("where") and feature ("what") information observable over short time scales (100s of ms to seconds) in both human infants and adults (reviewed by Leslie, Xu, Tremoulet & Scholl, 1998; Flombaum, Scholl & Santos, 2008). Subsequent investigation has shown that the "what" and "where" processing streams themselves diverge into substreams that process information about the features and motions of animate and inanimate objects, respectively, and has clarified both the functional architecture of visual processing in adults and its development in infants (reviewed by Gerhardstein, Schroff, Dickerson & Adler, 2009; Nassi & Callaway, 2009). The results of these studies are summarized in cartoon form in Fig. 1.

*****
Fig. 1 about here
*****

Both "what" and "where" pathways provide input to parahippocampal areas of medial-temporal cortex that construct integrated object and layout representations for hippocampal encoding into episodic memories (reviewed by Eichenbaum, Yonelinas & Ranganath, 2007; Zimmer & Ecker, 2010). The "where" pathway is also coupled directly to motor-planning components of posterior parietal cortex and, via the "mirroring" functionality of the motor-planning system, to prefrontal representations of goals associated with actions similar to observed actions and frontal (motor) representations of the muscular forces required to execute actions similar to observed actions (reviewed by Rizzolatti & Craighero, 2004; Cattaneo & Rizzolatti, 2009). It is this mirror functionality, which can be detected in infants from 9 months using electroencephalographic methods (reviewed by Marshall & Meltzoff, 2011) but which behavioral studies suggest is present from birth (Simion, Regolin & Bulf, 2008), that enables the inference from differential looking times to the representation of goals or intentions in studies such as that of Luo (2011). Posterior parietal systems also provide "how" and "why" information, representing both the motion followed by an object from its previous location and the actions responsible for that motion, to the hippocampus; as the motion followed by an object may have been unobserved, this "how" and "why" information may be synthesized by the action-planning system on the basis of previous episodic memories involving and categorical constraints applying to the object (reviewed by Fields, 2012a). The ability to deploy "how" and "why" information to re-identify objects across unobserved changes in location or features is evident in preschoolers (e.g. Gutheil, Gelman, Klein, Michos & Kelaita, 2008), but has not been tested directly in infants.

Within the context of a characterized functional architecture such as that sketched in Fig. 1, reproducible activations of identified components can be assigned a semantics and hence considered to be "representations" of the objects, features, motion patterns, events or actions that reproducibly activate them. For example, activations of cells of the lateral fusiform gyrus (LFG) that reproducibly respond specifically to human-like faces in a location- and context-independent way can be said to "represent" the property "has a human-like face" (e.g. Yovel & Kanwisher, 2004); similarly, activations of motion-responsive but not feature-responsive cells in the superior temporal sulcus (STS) that reproducibly respond to objects moving on horizontal linear trajectories in a context-independent way can be said to "represent" a horizontal linear trajectory in a location and scale-invariant way (e.g. Fields, 2011). The use of this sense of representation, which has its origin in general systems theory (e.g. Ashby, 1956), in the context of functional and cognitive neuroscience dates back at least to the early 1980s (e.g. Marr, 1982; Anderson, 1983; Rock, 1983; Pylyshyn, 1984). It is dependent on the functional roles of the components characterized as representational within the overall functional architecture, i.e. on the actual flow of information and control through the system (Fields, 1994). This dependence of the semantics on the architecture as a whole is often under-appreciated. Haith (1998), for example, complains that considering "the coding of information in neural networks" to be "representation" is too all-inclusive and hopes that "infant researchers must mean more than that" (p. 173) but neglects the constraints imposed by the architecture on *what* information any particular component can be considered to encode. Such constraints are clearly critical in any discussion of the innate encoding of a "principle" such as the principle of persistence.

A straw man

The words "principle" and "concept" call to mind a particular kind of reasoning: conscious, deliberative, language-based reasoning from a general principle to a specific conclusion. One can, for example, reason from the Principle of Conservation of Energy to the conclusion that a perpetual-motion machine can do no useful work, or from the Principle of the Excluded Middle to the conclusion that if a proposition *P* is not true, *P* must be false. Adults and adolescents are competent in this kind of language-based reasoning to varying degrees; pre-verbal infants are not. There is a common tendency, however, to conflate *representation* with *language-like representation*, a tendency exemplified by the notion that "symbolic" reasoning is implemented by a "language of thought" sharing the grammatical and semantic features of natural languages (Fodor, 1975; 2008). Haith (1998) appears to be committing this conflation when he remarks, "if we mean that infants can *re*-present events to themselves by calling them up from memory, to generate a schema or image that they reason about, create expectations and beliefs from, and make inferences about, then we are talking about something else, something that begins to sound like a symbolic representation" and then characterizes as an "absurd example" the idea that infants (or anyone) "must represent the environment to themselves during (a) blink" of 100 ms (all p. 173, emphasis in original). If "representation" is viewed in terms of specific activation as discussed above, infants (or anyone) clearly *do* represent the environment to themselves during a blink; indeed the "lingering sensory activations" favored by Haith as an explanation of object persistence implement just such a representation.

If the straw man of language-like representations is rejected, the question of how object persistence is implemented by the visuo-motor system can be clearly framed. Any object that is represented within short-term memory or longer-term episodic memory and then later re-identified as the *same thing* is clearly being represented as persistent, whether this representation is accurate – i.e. whether the re-identified object really is the same thing as the remembered object – or not. The representation of something as persistent in this operational sense of re-identifiable over time as the same thing that was seen previously is, moreover, entirely independent of any representation, linguistic or otherwise, of "persistence" as an abstract concept: there is no need to postulate that infants understand "object persistence" as a concept when they re-identify their mothers upon waking up in the morning. It is, moreover, independent of whether any memory of the object is consciously retrievable. All that is required, even over substantial gaps in observation, is that what Zimmer and Ecker (2010) call an "object token" – a persistent, individual-specific representation of a located, featured object – is encoded by the hippocampus. If this is the requirement for the representation of object persistence, then the criteria for constructing such a representation – the criteria specified by the persistence principle – must be implemented by the ventral "what" pathway, the dorsal "where" pathway, or their interactions. To the extent that these pathways are well-defined, the question of what representations they implement is an empirical question addressable with available data and further experiments.

Is the persistence principle implemented by categorization?

Objects, properties, events and actions are fundamental to the conceptual structures of typical adults. Objects *have* properties, *participate in* events, and if they are agents, *perform* actions. These fundamental ontological distinctions and their relationships develop early and are clearly evident in preschoolers (Karmiloff-Smith, 1995); indeed preschoolers with suitable experience demonstrate a more subtle ontology that acknowledges the possibility of agent-independent mechanical causation (e.g. Sobel, Yoachim, Gopnik, Meltzoff, & Blumenthal, 2007; Sobel & Buchanan, 2009). Interacting successfully with objects requires learning what their properties (typically) are, what kinds of events they can (typically) participate in, and what kinds of actions, if any, they can (typically) perform. That the latter two kinds of knowledge develop more slowly than the first is well documented (e.g. Karmiloff-Smith, 1995; Kelemen, 2004).

Knowledge of the typical properties, occurrence contexts and capabilities of objects is *categorical* knowledge. While the implementation of categorical knowledge, and in particular the relationship between categorical knowledge and the natural language that allows its verbal expression is far from fully characterized, it is clear that the what, where, how and why pathways contribute to object categorization by representing the static properties of observed objects, the spatiotemporal structures of events, and via mirror excitations, the typical goals of and effort required by actions not just as bottom-up sensory activations, but as resonant activation patterns that incorporate top-down, long-term-memory resident categorical knowledge (reviewed by Martin, 2007; Zimmer & Ecker, 2010; Fields, 2012a; Kiefer & Pulvermüller, 2012 from a cognitive neuroscience perspective; see e.g. Grossberg, 2013 for a computational modeling perspective). Identifying something as a horse or a baseball game, for example, requires satisfying categorical constraints on what *counts as* a horse or a baseball game; resonance between bottom-up and top-down activation patterns implements this constraint satisfaction.

It is natural, therefore, to ask whether the object categorization process implements the persistence principle, i.e. to ask whether the categorization process imposes the constraints of continuous existence, cohesiveness and property retention as categorical constraints on all members of the category “object.” Implementation of the persistence principle by the categorization process could be expected on the basis of the traditional role of an “object” devoid of any specializing features in philosophical discussions (Leibnizian monads are arguably such things), or the common practice of employing “physical object” as the featureless root node of a category hierarchy in software implementations of “upper-level” ontologies that support tasks such as natural-language processing (e.g. Sowa, 1984; Niles & Pease, 2001). If the human categorization process employed a well-defined category hierarchy, one might expect that the first test applied to any percept would be to determine whether it represented a “physical object”; such a test would provide a natural place to require continuous existence, cohesiveness and the retention of properties.

Human infants are sensitive to features such as geometric shape at or soon after birth, and are capable of category learning from exemplars as early as four months (reviewed by Rakison & Yermoleva, 2010). Subordination from entry-level categories such as “person” or “dog” to lower-level categories such as “boy” or “dachshund” has been observed in infants as young as 6 months (Quinn & Tanaka, 2007). By 9 months, object categorization is well established, although category boundaries do not necessarily correspond to adult category boundaries (reviewed by Mandler, 2004). Xu (1997) argued on the basis of earlier studies showing that 4 to 5 month-old infants are able distinguish one object from two when they were presented simultaneously but unable to distinguish the same objects on the basis of their properties when they were presented separately that infants possess a *sortal* category “physical object” prior to learning more specific sortal categories – i.e. categories supporting object individuation and counting – such as “dog” or “person.” If infants possess a sortal category “physical object” at birth, or as Xu (1997) suggests as their *first* sortal category, perhaps the process of recognizing instances of this category imposes the requirements for continuous existence, cohesiveness and retention of properties that compose Baillargeon's (2008) persistence principle.

Evaluating this potential implementation of the persistence principle by the categorization system requires examining both how information about objects is represented by the perceptual-motor system and how and when these representations become functionally available during infancy. The category learning and sequential object distinction studies discussed above, together with many other studies, establish that infants become attentive to the visible features of objects gradually, and in a predictable order, during the first year of life (reviewed by Baillargeon *et al.*, 2011; 2012). Needham, Dueker and Lockhead (2005), for example, found that four and a half month-old infants do not visually segregate displays of unfamiliar, stationary solid objects into distinct, bounded entities based on their shapes; they were unsurprised whether adjacent, differently-shaped solid objects moved as one entity or as two (see also Xu, 1999 for a discussion of similar experiments with visually more-complex objects). Given appropriate experience with other objects, similar to those displayed, presented as individually-movable items, however, the same infants segregated adjacent solid objects based on shape in the way children or adults would; such experience dependence is consistent with that observed by Luo (2011) in the moving-box experiments discussed earlier. These results suggest that even geometrical shape, one of the most basic features an object can have, is not automatically salient at four and a half months, but can become salient with suitable experience. Experiments employing occluders (e.g. Luo & Baillargeon, 2005) and containers (e.g. Wang, Baillargeon & Brueckner, 2004) of various sizes relative to the object occluded or contained have similarly shown that size, another quite basic feature, is not automatically salient to the youngest (2.5 - 3 month old) infants, and that sensitivity to size relationships typically develops at different ages for different relationships, again in an experience-dependent way. Surface color and color pattern become salient as persistent features of objects at even later ages. Taken together these results suggest that, with the prominent and ecologically-significant exceptions of human faces and hands, the visible features of objects are not salient to the youngest infants, and become salient only with suitable experience. Such gradual development of feature salience is consistent with the delayed development of specific sortal categories and hence of object individuation on the basis of even entry-level category membership as noted by Xu (1999).

Both the non-salience of visual feature information for the youngest infants and the gradual development of feature salience in older infants are consistent with the earlier development of the motion-processing dorsal visual stream relative to the feature-processing ventral visual stream in humans (Leslie *et al.*, 1998; Gerhardstein *et al.*, 2009). Even the youngest infants are highly sensitive to visual motion, and demonstrably employ motion-based cues to identify objects in visual scenes. While the visual cues indicative of object persistence at four months are not identical to those indicative of object persistence to older children or adults (Bremner, Johnson, Slater, Mason, Foster, Cheshire & Spring, 2005; Bremner, Johnson, Slater, Mason, Cheshire & Spring, 2007), persistent object detection during short (seconds) visual encounters is essentially feature-independent for infants as it is for older children or adults (reviewed by Flombaum, Scholl & Santos, 2008). In adults, high-level or superordinate object categorization (e.g. animal v/s non-animal) based on features and/or motion patterns requires on the order of 200 ms (Kiefer, 2001; Tkach, Reimer & Hatsopoulos, 2007; Macé, Joubert, Nespoulous & Fabre-Thorpe, 2009; Mukamel, Ekstrom, Kaplan, Iacoboni & Fried, 2010), considerably longer than the 50 ms required for visual short-term memory consolidation (Vogel, Woodman & Luck, 2006) and hence object-file (Treisman, 2006; Flombaum, Scholl & Santos, 2008) construction. In adults, therefore, features are additions to *pre-existing* representations of persistent objects, i.e. object files, constructed on the basis of trajectory information alone (Flombaum, Scholl & Santos, 2008; Fields, 2011). All available evidence suggests that the same is true for infants.

The rapidity of object file construction and subsequent feature binding and categorization compared to the roughly 270 ms required for consciousness of a visual display (Sergent, Baillet & Dehaene, 2005) rules out any role for conscious deliberation in determining whether something seen is an *object* under ordinary circumstances. Under ordinary circumstances, human beings see categorized objects; they do not see collections of features that they *infer* to be objects, or uncategorized objects for which they must consciously infer a category. They see these categorized objects, moreover, in the context of ongoing events. “Event files” (Hommel, 2004) binding objects to actions and even to emotional responses are constructed within 240 to 280 ms (Zmigrod & Hommel, 2010); hence “how” and “why” information are bound to object representations within the time required for consciousness (Fields, 2012a). It is, indeed, not unreasonable to suppose that coordinated activity across the network shown in Fig. 1 implements the “global workspace” (Baars, 1997; Dehaene & Naccache, 2001) of ongoing, conscious visual awareness, awareness that always represents objects as persistent.

The temporal precedence of object-file construction, however, renders the idea that a foundational concept or category of “persistent object” imposes the constraints of continuous existence, cohesion and property retention on observed objects functionally and hence theoretically irrelevant except as a post-hoc, adult construct. When a moving object with a recognizable trajectory is observed, or an object enters the visual field through movement of the head or the motion of something in the environment, the construction of an object file by the dorsal processing stream implicitly defines the object *as an object*, i.e. as a localized, possibly moving area within visual space to which features can be bound, within roughly 50 ms. Detected features are bound to this object file roughly 150 ms later, while "how" and "why" information is added and multiple object files combined to form an event file roughly 50 ms after that (Fields, 2011). Categorization implemented by the ventral stream plays no role in the 50 ms process of identifying something as an object by constructing an object file; categorization is a follow-on process of reconciling detected features with those encoded by available category representations to determine *what kind* of object is represented by the object file. The ventral stream may "know" that a human-like face, for example, is present in a scene due to activation in the LFG, and hence may "expect" that the face is part of a human being on the basis of categorical knowledge, but does not "know" that an *object* with a human-like face is present until after the face feature has been bound to an object file. Categorization of an observed object as a person, therefore, is implemented by dorsal-ventral representation binding, a process that happens well after the representation of what is being observed as an "object" by object-file construction. Supposing the existence of a root-level category "physical object," either as a purely descriptive category or as a sortal, clearly adds nothing to this picture. By definition, a root-level "physical object" has no specific and distinguishing features; "physical object" as a category must encompass all objects, whatever their features. Hence there are no features to be bound to an object file that mark it as a "physical object." Categorizing the content of an object file as a "physical object" at some point further downstream of feature binding, moreover, adds no information to that already provided by the bound object file itself: anything represented by an object file has already been determined to be cohesive and continuous, at least during the ongoing period of observation, and it has already been determined to have whatever features are bound to it. Thus while adult human beings in at least some cultures possess a verbalizable abstract concept "physical object" which serves as a sortal, this abstract concept need not correspond to any part of the object categorization process implemented by the ventral stream. Such a concept certainly can play no role for infants or children for whom such high-level abstractions are at least a decade away (e.g. Karmiloff-Smith, 1995). If the persistence principle is taken to be implemented by the ventral-stream object categorization system, therefore, it is theoretically vacuous.

It may be objected that while ventral-stream object categorization follows object-file construction, it is required in some cases to determine whether something represented as an "object" in the sense of having its own object file is indeed an object "in its own right" or is actually a feature of some larger object; for example, whether a black spot localized in visual space is an object or a feature on a leopard. Unmoving, always-present objects, in particular, can only be recognized as "objects" if their features are recognized, by the categorization process, as being features that are specified within some category as being features of an *object* as opposed to merely being localized features of the background. Even in these cases, however, persistence is conferred by the object file itself, not by the object categorization process: object categorization applies to, and applies only to, already existing object files that represent *something* as continuously existing within the period of observation. A "something" in the visual field is, moreover, never perceived as a "bare particular" without features; an object file alone is insufficient for visual awareness. Even in the experiments of Gao and Scholl (2010), which demonstrated that adults can identify moving objects that are visually indistinguishable from the surrounding background when not moving, features such as size and shape were reliably assigned when the objects were distinguishable from the background, i.e. when they were moving. As mentioned above, infants are sensitive to features such as object shape from birth (Rakison & Yermoleva, 2010); such features may in some or even most cases not be salient, e.g. as demonstrated by Needham, Dueker and Lockhead (2005), but they are not *invisible*. Experiments such as those discussed by Xu (1997) in which infants are tasked with distinguishing objects presented either together or separately are uninterpretable unless the infants are assumed to be consciously aware of the objects they are observing, and hence consciously aware of features that distinguish them from the background, even if the distinguishing features are not salient and are not remembered across trials. The observed ability of infants to distinguish objects presented together even though they could not distinguish them when presented separately can be explained by their inability to remember features across trials together with the structure of the event file. Infants see two objects, with features but otherwise uncategorized, if their awareness is driven by an event file that incorporate two object files, and see just one object, with features but otherwise uncategorized, if their awareness is driven by an event file that incorporates just one object file. The infants behave as if they employed "physical object" as a sortal category, but there is no need to postulate an implemented "physical object" categorization process to explain their behavior.

Object persistence as an outcome *of trajectory recognition*

It can be argued, of course, that glossing "initial knowledge"or "innate idea" in terms of categorization is an over-interpretation of Spelke's (1994) and Baillargeon's (2008) intentions regarding their "principles" of infant reasoning. Perhaps it is intended that these principles are implemented by the process of object-file construction, independently of whether the "object" represented by an object file is later determined to merely be a feature, such as a spot on a leopard or a light on the left elbow of a point-light walker. Object-file construction on the basis of trajectory information is a computational procedure, and it is regularly characterized as employing "principles" such as spatiotemporal priority (e.g. Flombaum, Scholl & Santos, 2008). Perhaps the cohesion and continuity components of the persistence principle should be viewed as implemented by dorsal-stream object-file construction, while the property retention component should be viewed as implemented as a later, secondary process by ventral-stream feature detection and object categorization.

The first thing to note regarding this alternative reading of the persistence principle is that it is only sometimes descriptive and is never explanatory. Cohesiveness cannot be a *post-hoc* criterion of objecthood; object files are not constructed and then checked somehow for cohesiveness. The representation of cohesive objects by object files is, rather, an outcome of the criteria employed to construct object files, criteria that have only the motion-vector representation constructed by MT (Nassi & Callaway, 2009) to work with. Continuity, similarly, cannot be a *post-hoc* criterion for trajectory recognition; trajectory continuity is similarly an outcome of object-file instantiation criteria defined at the MT level of representational complexity. Indeed, there is no process by which an object file "become(s) subject to the persistence principle" (Baillargeon, 2008, p. 3); if something is represented by an object file, it is represented as persistent. If the persistence principle is to be explanatory, therefore, it must explain the construction of object files, with the representations available at the level of MT as input. The simplest model of object-file construction that explains the phenomenology while remaining consistent with what is known about the functional architecture of the dorsal stream posits the experience-dependent development of specific, post-MT recognition networks for each of the curvilinear trajectories that indicate continuity and hence objecthood to adults (Fields, 2011). It is well-known that not all curvilinear trajectories indicate continuity to human observers, and that the trajectories that do so are exquisitely sensitive to the presence of occluders. An "object" that "moves" along a "trajectory" that is not recognized as a trajectory by the post-MT visuo-motor system is not perceived as a persistent object, and hence is not perceived as moving, consistent with the many studies that demonstrate perceived object discontinuity when trajectory-recognition criteria are violated (Flombaum, Scholl & Santos, 2008; Fields, 2011). In this model, persistence is an outcome of downstream network sensitivity to some but not all velocity-segment maps produced by MT, while spatiotemporal priority is an outcome of the earlier maturation of the dorsal stream. Postulating a "principle of persistence" adds nothing to this description of how the process works; conversely, it explains nothing about how the process works. There may be many "objects" that even adult human beings do not see as moving, do not construct object files for and hence do not recognize as objects, because our post-MT visuo-motor systems cannot recognize and respond to their trajectories. As pointed out in Fields (2011), for example, the "fixed stars" were seen as *features* by medieval cosmologists.

The role of trajectory recognition in determining not only what counts as an object but what kind of object it counts as is illustrated particularly clearly by visual responses to point-light walkers (reviewed by Puce & Perrett, 2003; Flombaum, Scholl & Santos, 2008). Infants respond differentially to point-light walker displays even at birth (Simion, Regolin & Bulf, 2008); by 6 months there are sensitive to a point-like walker's direction of motion (Kuhlmeier, Troje & Lee, 2010). Adult recognition of point-light walkers as coherently moving objects requires less than 100 ms (Pavlova, Birbaumer & Sokolov, 2006), indicating that a point-light walker is represented as an object, by an object file, prior to feature binding and hence prior to any input from ventral-stream object categorization. A point-light walker display is only cohesive to the extent that the relative distances between the point lights remain within certain scale-independent bounds. Similarly, the apparent motion of a point-light walker display is only continuous to the extent that the relative motions of the point lights maintain a particular dynamically- but not statically-specifiable organization. The relative distances between lights and their relative motion together define a collective trajectory that must be recognized to initiate construction of an object file representing the point light walker. The cohesion and continuity of the walker are *outcomes* of the recognition of this collective trajectory, not criteria applied the representation of the walker following recognition. One cannot explain the recognition of a walker by pointing to cohesion and continuity; one can only explain it by pointing to the collective trajectory itself.

The second thing to note about this alternative reading of the persistence principle is that a principled requirement that objects "retain their individual properties" (Baillargeon, 2008) effectively enforces "Leibniz's Law" of the identity of indiscernibles as a criterion of object identity over time: if two objects seen at two different times have identical features, then they must be identical, i.e. they must be the *same thing*, while if their features are not identical, they cannot be the same thing. Treating indiscernibility as a necessary criterion of object identity over time clearly fails for any objects that in fact change their properties through time, as living things, natural objects and even artifacts inevitably do. Treating indiscernibility as a sufficient criterion of object identity over time fails in any environment containing multiple indiscernible "copies" of any particular object. The contemporary world of technological artifacts is replete with such copies; even the environments of hunter-gatherers are replete with effectively indiscernible forest trees, herd animals, and similar natural "copies." Hence while Hollingworth and Franconeri (2009), for example, claimed that "under normal circumstances, a representation of the surface features of an object will be sufficient to distinguish it from other visible objects" (p. 115), under many common circumstances they are not. Any application of the persistence principle to properties, therefore, requires caveats and exceptions: properties of an object that can change with time (sometimes called "inessential properties" in the philosophical literature) cannot be used as criteria of object identity, and objects that are known to exist in multiple copies cannot be identified by even their essential properties. The problem with such caveats and exceptions is evident: one cannot determine what properties of an object might be essential unless one can re-identify it over time, and one cannot determine that multiple copies of an object exist unless one can distinguish and compare them. Hence characterizing the object-identification system as a procedural implementation of the persistence principle as applied to properties is once again only partially descriptive at best; *how* the system implements the principle, and how it learns as well as handles the various required exceptions and caveats remain to be explained. Indeed, even viewing property persistence as a *principle* of object identification is misconceived; what humans regard as objects – what humans instantiate object files in response to – only have some visually-evident properties that persist, such properties persist only some of the time, and the association of such properties with an object only takes place following object-file construction.

In summary, claiming that either infants or adults employ a "persistence principle" to segregate objects from their backgrounds or to re-identify objects over time appears to contribute nothing to understanding either why or how human beings accomplish these perceptual and cognitive feats. *Why* humans or other animals would segregate the perceptible world into objects that could be re-identified over time seems tolerably clear on evolutionary grounds; indeed evolution cannot even be conceived as a process without assuming the existence of individual temporally-persistent organisms. *How* they do it, and in the case of infants *if* they do it, is another matter. Gerhardstein *et al*. (2009) counter the nativism of Spelke or Baillargeon with the claim that "the earliest skills available to the infant are those clearly adaptive for survival or those that regulate early exposure to visual content, thereby supplying the system with meaningful visual experience that will *eventually* facilitate object perception " (p. 80; italics added). If the assumption that infants employ a persistence principle is set aside as theoretically unhelpful, experiments such as those of Luo (2011) or Needham, Dueker and Lockhead (2005) cry out for more detailed interpretations. What do 3-month-old infants see, for example, when we show them a box moving toward a cone (Luo, 2011)? Do they see a *box*? Or do they just see an *object*? When we then show them a box moving toward a cylinder, do they really see the *same* box moving? Or do they just see a slightly different scene involving a moving object, one just different enough to attract their attention? What do four and a half month-old infants see when we show them an arrangement of unmoving shapes (Needham, Dueker & Lockhead, 2005). Do they even see *objects* before the shapes move, or do they just see features of the otherwise-neutral background? Addressing these questions requires the consideration of implementation details that are papered over by the assumption of an in-born principle of object persistence. Viewing object persistence as an outcome – as an effect to be explained – opens these questions up for experimental investigation.

The computational task of object re-identification

While saying that the perception of object persistence is guaranteed by a "principle" has no theoretical force, human object perception nonetheless displays a quite significant "persistence effect": human beings tend to perceive objects as persistent over time, and human behavior would be barely comprehensible otherwise. This persistence effect demands explanation. Postulating specific trajectory-recognition networks in visuo-motor cortex provides a partial, empirically-testable explanation of the perception of object persistence in scenes lasting at least a few seconds. Human beings do not, however, follow the trajectories of familiar objects at all times; indeed at any given time, most of the objects familiar to any given person are out of sight. Hence what demands explanation is the ability of human beings to re-identify objects as familiar – as previously-encountered individuals – across substantial gaps in observation. How, for example, do infants re-identify their mothers when they wake up in the morning? Such questions are significant because re-identification can fail, for example in autism (Fields, 2012b).

It is worth emphasizing that the visual object re-identification problem, like the visual object segregation problem, is computationally intractable. It is well known that the information in a static visual array is insufficient for object segmentation (e.g. Gregory, 1970; Marr, 1982); this is one reason why the ability to grasp and manipulate objects is so important for infants. The object re-identification problem exhibits a similar lack of sufficient information; indeed the object re-identification problem can be viewed as a special case of the computationally-intractable "frame problem," the problem of determining what has not changed as a result of an action (McCarthy & Hayes, 1969). In the case of object re-identification, what must be determined not to have changed is an object's *identity*; the insufficient information in this case is the in-principle unobtainable information about what possible causal processes might have occurred during a period of non-observation (Fields, 2013). Humans cope with computational intractability by deploying heuristics; hence the first question posed by the "persistence effect" is the question of what heuristics humans actually employ to re-identify objects. Heuristics are *not* principles. Heuristics provide neither necessary nor sufficient conditions; they are rough and ready solution methods that work fast enough, often enough and well enough not to be abandoned, but they provide no guarantee of success. They are good guesses about how the world works, not knowledge.

The exigencies of experimental design naturally divide studies of object re-identification into three temporal domains: a few seconds, a few minutes, and longer periods. Research in the few-second domain was pioneered by Burke (1952); Flombaum, Scholl and Santos (2008) provide a thorough review. Such studies provide extensive evidence for spatiotemporal priority in the identification of persistent objects in short perceptual encounters; some designs (e.g. Hollingworth & Franconeri, 2009) reveal featural contributions in this time-frame, but most do not. Research in the few-minute domain is exemplified by the experiments of Luo (2011) or Needham, Dueker and Lockhead (2005); the studies reviewed by Baillargeon *et al*. (2011; 2012) fall into this domain, as do some studies of object re-identification by older children (e.g. Gutheil *et al*., 2008). Object re-identification over longer periods is often investigated by non-experimental means, including natural-historic observation, self-reports, and the use of thought experiments (reviewed by Rips, Blok & Newman, 2006; Scholl, 2007; Xu, 2007). Work in these latter two temporal domains often explicitly targets situations in which Leibniz's Law fails; the work of Gutheil *et al.* (2008), for example, shows that four- and five-year olds can employ information about the causal histories of objects to differentiate between otherwise-indiscernible "copies" (in this case, identical dolls or plush toys).

A consistent result of re-identification studies in both the few-minute and longer time domains, whether in the laboratory or in the everyday world, is that human beings cope with both featural change and the presence of featurally-identical or nearly-identical distractors by appealing to either individual or categorical constraints on the possible causal histories of objects between observations. In a study by Hood and Bloom (2008), for example, preschoolers were not only able to distinguish, but showed a decided preference for, an item of their own as opposed to an exact "copy" purportedly made by a machine belonging to the experimenters. A world in which *some* things are known to undergo featural changes with time is, moreover, a world in which *all* things could, at least in principle, undergo such changes; learned categorical criteria pertaining to featural change are required to specify whether and what kinds of featural changes can occur to what kinds of objects (e.g. Xu, 2007). Similarly, a world in which some objects can be perfectly duplicated is a world in which all objects potentially could be; learned categorical criteria pertaining to duplication are required to determine the probabilities of or conditions under which duplication might be expected in any particular case. An evaluation of the plausibility, given all relevant individual and categorical criteria, of any object seen now being the very same thing as some object seen previously is, therefore, to be expected in every case of object re-identification across a gap in observation long enough for significant causal processes to act. Such evaluations of causal plausibility are necessarily heuristic; the categorical constraints on motion and change that they access are experience-dependent, fallible and subject to revision. As discussed above, the rapidity of such evaluations, as well as the results of adult functional imaging studies of object re-identification in association with episodic-memory recall (reviewed by Moscovitch, 2008; Ranganath, 2010), suggest that they are implemented pre-consciously by the parietal-lobe pre-motor processes that implement action planning (Fields, 2012a).

If the functional model of Fields (2012a) is correct, the ability to re-identify objects, and hence the ability to regard individual objects as persisting through time, develops in tandem with the ability to represent causal processes. Seeing an object move or be moved indicates persistent objecthood, for at least the period during which the object is observed. Seeing a second, featurally-similar object move or be moved, as in the additional experience phase of the experiments of Needham, Dueker and Lockhead (2005), also conveys objecthood. What remains unclear, within this model, is the extent to which objects can be segregated from their local backgrounds in the absence of both motion and categorical knowledge. Experiments specifically designed to probe the ability of infants to distinguish objects from background features in the absence of within-scene motions, including the motions of screens or curtains that cause objects to "pop" into visual space, would address this question. For example, the classic experiment of Baillargeon, Spelke and Wasserman (1985) in which an occluder moved so as to rest on a previously-visible block could be modified by visibly removing the block after a familiarization trial to reveal a background feature of the same visual shape, size, color and shading. Such a modification would allow a direct test, using looking time as an assay, of whether infants require actual 3-dimensionality as a criterion for object segregation. Whether object segregation is preserved across brief gaps in observation for static as well as moving objects could be tested by modifying other standard designs, such as that of Xu and Carey (1996), by replacing previously-manipulated 3-dimensional objects with visually-similar background features and testing infants' responses to "impossible events" involving each.

Conclusion

Two things have been argued here. First, the notion that either infants or adults employ a "principle of persistence" to reason about objects is theoretically unhelpful; the persistence principle only partially redescribes, without explaining or even shedding any light on, the ability of human beings to segregate and re-identify objects. This should not be too surprising. As pointed out by Cummins (1983), scientific "principles" are often not explanatory: the Principle of Conservation of Energy does not explain energy conservation, while the Principle of the Excluded Middle does not explain why sentences are either true or false, as opposed to something in between. The fact that humans perceive objects as persistent through time is, rather, an outcome to be explained, on a par with the Stroop Effect or any of the myriad other "effects" of cognitive psychology. Second, learned individual and categorical constraints on how object features can change and whether objects can be duplicated are significant contributors to object re-identification and hence to the inference of object persistence over substantial gaps in observation. Such learned constraints are necessary to support heuristic "best guesses" about whether an object seen now is the same individual thing as an object seen previously. *Guessing* is necessary because the problem of object re-identification is computationally intractable; there are no (finite) principles by which correct object re-identification can be guaranteed.

Understanding how infants re-identify objects requires both understanding how they make constraint-driven heuristic guesses about object identity, and understanding how this guessing process is implemented by the developing neurocognitive system.

If the persistence principle is rejected, the results of many if not most experiments probing infant abilities to visually segregate objects and interpret their motion over time periods longer than a few seconds require rethinking. Relatively straightforward modifications of classic experimental designs would enable direct tests of whether infants can segregate or re-identify objects in the absence of within-scene motion. Such tests would demonstrate the extent to which any “principle of persistence” can be taken for granted as even a partial description of infant cognition about objects.

Acknowledgments

Thanks to two anonymous referees for helpful comments on an earlier version of this paper.

References

Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.

Ashby, W. R. (1956). *An Introduction to Cybernetics*. London: Chapman & Hall.

Baars, B. J. (1997). In the theater of consciousness: Global workspace theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies 4*, 292-309.

Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants’ physical reasoning. *Perspectives on Psychological Science 3*, 2-13.

Baillargeon, R., Spelke, E. S. & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition 20*, 191-208.

Baillargeon, R., Li, J., Gertner, Y. & Wu, D. (2011). How Do Infants reason about physical events? In U. Goswami (Ed.) *The Wiley-Blackwell Handbook of Child Cognitive Development, 2nd Ed*. Oxford: Blackwell (11-48).

Baillargeon, R., Stavans, M., Wu, D., Gertner, Y., Setoh, P., Kittredge, A. K. & Bernard, A. (2012). Object individuation and physical reasoning in infancy: An integrative account. *Language Learning and Development 8*, 4-46.

Bremner, J. G., Johnson, S. P., Slater, A., Mason, U., Foster, K., Cheshire, A. & Spring, J. (2005). Conditions for young infants' perception of object trajectories. *Child Development 76*, 1029-1043.

Bremner, J. G., Johnson, S. P., Slater, A., Mason, U., Cheshire, A. & Spring, J. (2007). Conditions for young infants' failure to perceive trajectory continuity. *Developmental Science 10*, 613-624.

Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology 4*, 121–138.

Cattaneo, L. & Rizzolatti, G. (2009). The mirror neuron system. *Archives of Neurology 66*, 557-560.

Cummins, R. (1983). *The Nature of Psychological Explanation*. Cambridge, MA: MIT.

Dehaene, S. & Naccache, L. (2001). Toward a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition 79*, 1-37.

Eichenbaum, H., Yonelinas, A. R. & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience 30*, 123-152.

Fields, C. (1994). Real machines and virtual intentionality: An experimentalist takes on the problem of representational contents. In E. Dietrich (Ed.) *Thinking Computers and Virtual Persons*. New York: Academic Press (pp. 71-90).

Fields, C. (2011). Trajectory recognition as the basis for object individuation: A functional model of object file instantiation and object-token encoding. *Frontiers in Psychology – Perception Science 2*, 49 (doi: 10.3389/ fpsyg.2011.00049).

Fields, C. (2012a). The very same thing: Extending the object token concept to incorporate causal constraints on individual identity. *Advances in Cognitive Psychology 8*, 234-247.

Fields, C. (2012b). Do autism spectrum disorders involve a generalized object categorization and identification dysfunction? *Medical Hypotheses 79*, 344-351.

Fields, C. (2013). How humans solve the frame problem. *Journal of Experimental and Theoretical Artificial Intelligence* in press (doi: 10.1080/0952813X.2012.741624).

Flombaum, J. I., Scholl, B. J. & Santos, L. R. (2008). Spatiotemporal priority as a fundamental principle of object persistence. In B. Hood & L. Santos (Eds) *The Origins of Object Knowledge.* Oxford: Oxford University Press (pp. 135-164).

Fodor, J. A. (1975). *The Language of Thought.* New York: Crowell.

Fodor, J. A. (2008). *The Language of Thought Revisited*. Oxford: Oxford University Press.

Gao, T. & Scholl, B. J. (2010). Are objects required for object files? Roles of segmentation and spatiotemporal continuity in computing object persistence. *Visual Cognition 18*, 82-109.

Gerhardstein, P., Schroff, G., Dickerson, K. & Adler, S. A. (2009). The development of object recognition through infancy. In B. C. Glenyn and R. P. Zini (Eds) *New Directions in Developmental Psychobiology.* Hauppauge: Nova Science Publishers (pp. 79–115).

Gredebäck, G. & von Hofsten, C. (2007). Taking an action perspective on infants' object representations. *Progress in Brain Research 164*, 265-282.

Gregory, R. (1970). *The Intelligent Eye*. New York: McGraw-Hill.

Grossberg, S. (2013). Adaptive Resonance Theory: How a brain learns to consciously attend, learn, and recognize a changing world. *Neural Networks 37*, 1-47.

Gutheil, G., Gelman, S. A., Klein, E., Michos, K. & Kelaita, K. (2008). Preschoolers' use of spatiotemporal history, appearance, and proper name in determining individual identity. *Cognition 107*, 366-380.

Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behavior & Development 21*, 167-179.

Hollingworth, A. & Franconeri, S. L. (2009). Object correspondence across brief occlusion is established on the basis of both spatiotemporal and surface feature cues. *Cognition 113*, 150–166.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences 8*, 494-500.

Hood, B. M. & Bloom, P. (2008). Children prefer certain individuals over perfect duplicates. *Cognition 106*, 455-462.

Karmiloff-Smith, A. (1995). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT Press.

Kelemen, D. (2004). Are children “intuitive theists”? Reasoning about purpose and design in nature. *Psychological Science 15*, 295-301.

Kiefer, M. (2001). Perceptual and semantic sources of category-specific effects in object categorization: Event-related potentials during picture and word categorization. *Memory & Cognition 29*, 100-116.

Kiefer, M. & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex 48*, 805-825.

Kuhlmeier, V. A., Troje, N. F. & Lee, V. (2010). Young infants detect the direction of biological motion in point-light displays. *Infancy 15*, 83-93.

Leslie, A. M., Xu, F., Tremoulet, P. D. & Scholl, B. J. (1998). Indexing and the object concept: Developing 'what' and 'where' systems. *Trends in Cognitive Sciences 2*, 10-18.

Luo, Y. (2011). Three-month-old infants attribute goals to a non-human agent. *Developmental Science 14*, 453-460.

Luo, Y. & Baillargeon, R. (2005). When the ordinary seems unexpected: Evidence for incremental physical knowledge in young infants. *Cognition 95*, 297–328.

Macé, M. J.-M., Joubert, O. R., Nespoulous, J. L. & Fabre-Thorpe, M. (2009). The time-course of visual categorizations: You spot the animal faster than the bird. *PloS ONE 4*, e5927.

Mandler, J. M. (2004). Thought before language. *Trends in Cognitive Science 8*, 508-513.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Marshall, P. J. & Meltzoff, A. N. (2011). Neural mirroring systems: Exploring the EEG mu rhythm in human infancy. *Developmental Cognitive Neuroscience 1*, 110-123.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology 58*, 25-45.

Maunsell, J. H. R. & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Reviews of Neuroscience 10*, 363-401.

McCarthy, J. & Hayes, P. J. (1969). Some philosophical problems from the standpoint of artificial intelligence. In: D. Michie & B. Meltzer (Eds.) *Machine Intelligence 4.* Edinburgh: Edinburgh University Press (pp. 463-502).

Moscovitch, M. (2008). The hippocampus as a "stupid", domain-specific module: Implications for theories of recent and remote memory, and of imagination. *Canadian Journal of Experimental Psychology 62*, 62-79.

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M. & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Current Biology 20*, 750-756.

Nassi, J. J. & Callaway, E. M. (2009). Parallel processing strategies of the primate visual system. *Nature Reviews Neuroscience 10*, 360-372.

Niles, I. & Pease, A. (2001). Towards a standard upper ontology. In C. Welty & B. Smith (Eds) *International Conference Formal Ontology in Information Systems (FOIS-2001)*. New York: ACM (pp. 2-9).

Pavlova, M., Birbaumer, N. & Sokolov, A. (2006). Attentional modulation of cortical neuromagnetic gamma response to biological movement. *Cerebral Cortex 16*, 321-327.

Piaget, J. (1954). *The Construction of Reality in the Child*. New York: Basic Books.

Puce, A. & Perrett, D. (2003). Electrophysiology and brain imaging of biological motion. *Philosophical Transactions of the Royal Society of London B 358*, 435-445.

Pylyshyn, Z. W. (1984). *Computation and Cognition*. Cambridge, MA: MIT Press.

Quinn, P. C. & Tanaka, J. W. (2007). Early development of perceptual expertise: Within-basic-level categorization experience facilitates the formation of subordinate-level category representations in 6- to 7-month-old infants. *Memory & Cognition 35*, 1422-1431.

Rakison, D. H. & Yermoleva, Y. (2010). Infant categorization. *Wiley Interdisciplinary Review of*

Cognitive Science 1, 894-505.

Ranganath, C. (2010). A unified framework for the functional organization of the medial temporal lobes and the phenomenology of episodic memory. *Hippocampus 20*, 1263-1290.

Rips, L., Blok, S. & Newman, G. (2006). Tracing the identity of objects. *Psychological Review 133*, 1-30.

Rizzolatti, G. & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience 27*, 169-192.

Rock, I. (1983). *The Logic of Perception*. New York: Macmillan.

Scholl, B. J. (2007). Object persistence in philosophy and psychology. *Mind & Language 22*, 563-591.

Simion, F., Regolin, L., and Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences U.S.A. 105*, 809-813.

Sobel, D., Yoachim, C., Gopnik, A., Meltzoff, A., & Blumenthal, E. (2007). The blicket within: Preschooler's inferences about insides and causes. *Journal of Cognitive Development, 8*, 159-182.

Sobel, D. M. & Buchanan, D. W. (2009). Bridging the gap: Causality-at-a-distance in children's categorization and inferences about internal properties. *Cognitive Development 24*, 274-283.

Sowa, J. F (1984). *Conceptual Structures – Information Processing in Mind and Machine*. Reading, MA: Addison-Wesley.

Spelke, E. S. (1984). Preferential looking methods as tools for the study of cognition in infancy. In G. Gottlieb & N. A. Krasnegor (Eds) *Measurement of Audition and Vision in the First Year of Postnatal Life*. Westport, CT: Ablex (pp. 323-363).

Spelke, E. S. (1994). Initial knowledge: Six suggestions. *Cognition 50*, 431-445.

Spelke, E. S. (1998). Nativism, empiricism and the origins of knowledge. *Infant Behavior & Development 21*, 181-200.

Tkach, D., Reimer, J. & Hatsopoulos, N. G. (2007). Congruent activity during action and action observation in motor cortex. *Journal of Neuroscience 27*, 13241-13250 .

Treisman, A. (2006). Object tokens, binding and visual memory. In H. D. Zimmer, A. Mecklinger & U. Lindenberger (Eds) *Handbook of Binding and Memory: Perspectives from Cognitive Neuroscience*. Oxford: Oxford University Press (pp. 315-338).

Van Essen, D. C. & Maunsell, J. H. R. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends in Neurosciences 6*, 370-375.

Vogel, E. K., Woodman, G. F. & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance 32*, 1436-1451.

Wang, S.-H., Baillargeon, R. & Brueckner, L. (2004). Young infants' reasoning about hidden objects: Evidence from violation-of-expectation tasks with test trials only. *Cognition 93*, 167–198.

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 Psychologia 102*, 113-136.

Xu, F. (2007). Sortal concepts, object individuation, and language. *Trends in Cognitive Sciences 11*, 400-406.

Xu, F. & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology 30*, 111-153.

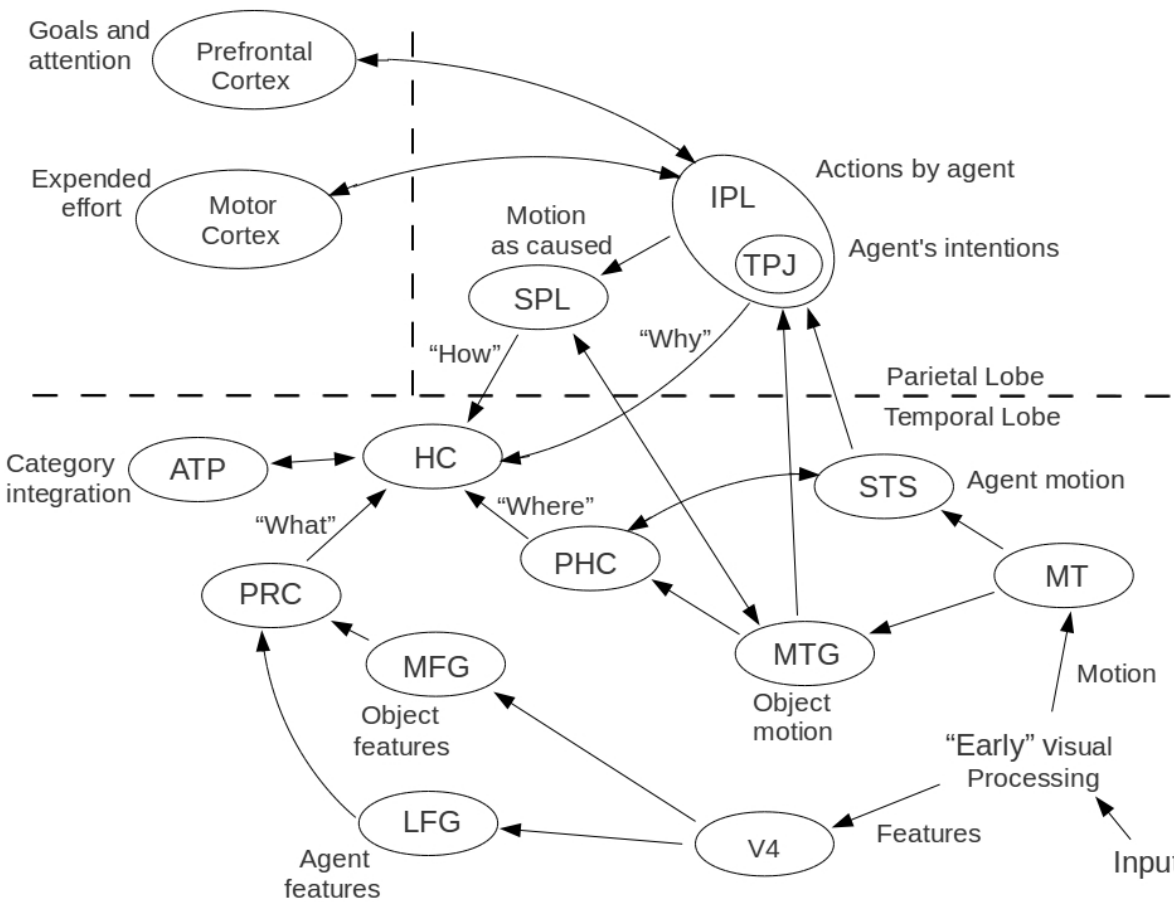
Yovel, G. & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron 44*, 889-898.

Zimmer, H. D. & Ecker, U. K. D. (2010). Remembering perceptual features unequally bound in object and episodic tokens: Neural mechanisms and their electrophysiological correlates. *Neuroscience and Biobehavioral Reviews 34*, 1066-1079.

Zmigrod, S. & Hommel, B. (2010). Temporal dynamics of unimodal and multimodal feature binding. *Attention, Perception, and Psychophysics 72*, 142-152.

Figure Caption

Fig. 1: Functional organization of visual-processing components of the human visuo-motor system. Abbreviations for functionally-characterized anatomical areas are as follows: V4, visual area 4; MT, medial temporal area; MTG, medial temporal gyrus; STS, superior temporal sulcus; PHC, parahippocampal cortex; LFG, lateral fusiform gyrus; MFG, medial fusiform gyrus; PRC, perirhinal cortex; ATP, anterior temporal pole; HC, hippocampus; SPL, superior parietal lobule; IPL, inferior parietal lobule; TPJ, temporal-parietal junction. Data are from Martin (2007), Cattaneo and Rizzolatti (2009), Gerhardstein *et al*. (2009), Nassi and Callaway (2009), Ranganath (2010), Zimmer and Ecker (2010) and Kiefer and Pulvermüller (2012). Figure adapted from Fields (2012b), Fig. 2.


Goals and attention
Prefrontal Cortex
Expended effort
Motor Cortex
Actions by agent
IPL
Motion as caused
TPJ
Agent's intentions
SPL
"How"
"Why"
Parietal Lobe
Temporal Lobe
Category integration
ATP
HC
STS
Agent motion
"What"
"Where"
PHC
MT
PRC
MTG
MFG
Motion
Object features
Object motion
"Early" Visual Processing
LFG
V4
Features
Input
Agent features