Visual re-identification of individual objects: A core problem for organisms and AI

Chris Fields

243 West Spain Street
Sonoma, CA 95476, USA

fieldsres@gmail.com

Abstract:

Two open questions about the visual re-identification of individual objects over extended time periods are briefly reviewed: 1) How much *a priori* information about the nature of objects, identity and time is required to support robust individual object re-identification abilities? and 2) How do epistemic feelings, such as the feeling of familiarity, contribute both to object re-identification and to the perception of opportunities and risks associated with individual objects and their affordances? The ongoing interplay between experiments that can be carried out with human subjects and experiments made possible with robotic systems is examined. It is suggested that developmental robotics, including virtual-reality simulations of robot-environment interactions, may provide the best route to understanding both the implementation of epistemic feelings in humans and their functional contribution to the identification of persistent individual objects.

Keywords: Causal reasoning; Developmental robotics; Epistemic feelings; Individual differences; Intrinsic motivation; Object persistence; Vision

Introduction

In the environments of many kinds of organisms, successful behavior requires an ability to reliably re-identify particular individual objects, for example particular individual conspecifics that may be allies or rivals, over periods of time extending from a few seconds up to many years. The human social environment takes this requirement to an extreme: a typical job, for example, may require categorizing thousands of objects using a multi-level class hierarchy and recognizing hundreds of these objects, including manufactured objects such as automobiles or computers as well as dozens of other human beings, as persistent individual entities with extended, detailed and highly task-relevant individual histories. Many artifacts also face this problem of re-identifying specific individual objects across various elapsed times. Computer operating systems, for example, must re-identify individual users, even if they do so solely by associating a user name and password with an identifier for a currently-active device. Scene-surveillance systems face the far more difficult problem (e.g. Satta, 2013) of tracking individual human beings as they traverse the fields of view of multiple cameras. Mobile, autonomous robots operating together in open environments may be expected to face object-individuation and individual object re-identification challenges not unlike those faced by humans.

Both the categorization of objects by type and the re-identification of particular individual objects over extended times have primarily been studied in the visual domain. Decades of detailed experimental work have demonstrated the critical role of motion within a visual scene for both human figure-ground distinction and object categorization (reviewed by Flombaum, Scholl & Santos, 2008; Fields, 2011a). Detecting motion requires detecting a persistent object that is moving; hence motion detection requires an inference of individual object identity over within-scene time intervals ranging from hundreds of milliseconds up to a few minutes. In the human visual system, such object identity inferences are typically executed in less than 100 ms, and construct the initial object representation – the "object file" (Kahneman & Treisman, 1984; reviewed by Treisman, 2006) – onto which feature identification and categorization are subsequently layered. On the longer end of the temporal scale, human beings are often able to re-identify specific individual objects – such as old friends or even former possessions – that have not been seen for years or even decades. This ability clearly involves matching features of the currently-observed object to those of individual-specific object tokens maintained in episodic memory (reviewed by Zimmer & Ecker, 2010). However, the robustness of human long-term individual re-identification abilities in the face of feature change or close featural competitors indicates that re-identification of observed objects as specific individuals involves causal reasoning that is highly sensitive to both general and context-specific knowledge (reviewed by Eichenbaum, Yonelinas & Ranganath, 2007; Fields, 2012a).

As philosophical paradoxes from the Ship of Theseus onward demonstrate, the concept of "identity" or "the very same thing" invoked in individual object identity judgments across time spans long enough to require the retrieval of object tokens from episodic memory is logically problematic; some notion of "causal continuation" as a criterion of long-term object identity is generally proposed to resolve these logical difficulties (reviewed by Scholl, 2007). In the case of actions or sequences of actions that are extended over such long periods of time, one must determine which of the objects that are present after the action has been completed are causal continuations of objects that were present before the action was initiated; similarly, one must determine which post-action objects are *not* causal continuations of pre-action objects and are therefore novel. This problem of determining which post-action objects are causal continuations of pre-action objects is, at least within the human cognitive architecture, equivalent to the frame problem (Fields, 2013a). The frame problem is computationally intractable and admits only heuristic solutions; hence only heuristic solutions of the long-term post-action individual object re-identification problem are computationally feasible.

The present review focuses on two questions about the individual object re-identification abilities displayed by humans and needed by autonomous robots in open task environments:

1. How much *a priori* information about the nature of objects, identity and time is required to support robust individual object re-identification abilities, or conversely, robust novel individual detection?

2. How do epistemic feelings, such as feelings of familiarity or uncertainty, contribute both to individual object re-identification and to the perception of opportunities and risks associated with individual objects and their affordances?

Behavioral and neurocognitive approaches to these questions are increasingly being complemented by modeling approaches using either physically-implemented robotic platforms or virtual-reality simulations of robot-environment interactions. As both of these questions are naturally asked from a developmental perspective in which abilities to re-identify individual objects are compared across the lifespan, work in the emerging field of developmental robotics (reviewed by Asada *et al*., 2009; Cangelosi and Schlesinger, 2015) is particularly relevant in this context.

The next section, “Motion detection, categorization and causal reasoning in individual object re-identification” focuses on the questions of what information is needed to re-identify individual objects and how much of it must be provided *a priori*. It begins by reviewing the neurocognitive processes by which human beings visually identify and re-identify individual objects over short elapsed times and how these processes develop from infancy to adulthood. It then discusses developmental robotic models of these processes and their utility as experimental surrogates, particularly for experiments that cannot ethically be carried out with human subjects. The re-identification of individual objects over longer elapsed times during which significant changes in object location or features may have occurred is then considered. In such cases, a plausible causal history connecting a remembered individual object to the currently-perceived object must be constructed (Fields, 2012a). How human subjects construct such causal histories is not well understood; robotic systems facing similar problems of individual object re-identification across extended gaps in observation may help shed light on this causal-reasoning process.

In humans, a distinct feeling that a particular object is familiar often both precedes and motivates the re-identification of that object as a previously-encountered individual. The third section, “Epistemic feelings as intrinsic motivators for individual object re-identification” reviews experiments in both human subjects and robots that indicate a broad role for familiarity, uncertainty, curiosity and other epistemic feelings in problem solving in general and individual object re-identification in particular. The roles of feelings of knowledge or certainty in terminating problem-solving or re-identification processes and marking solutions as such are also considered. Here too, developmental robotics may provide a route to understanding questions about human development and cognitive diversity that are otherwise experimentally intractable. The paper concludes that much remains to be done to understand individual object re-identification over both short and long timespans and lists some open questions.

*****
Box 1 about here
*****

Motion detection, categorization and causal reasoning in individual object re-identification

In humans, visual object identification begins with the instantiation of an object file in visual short-term memory by the position and motion sensitive dorsal visual processing stream. Object-file instantiation is highly dependent on the trajectory in visual space of an observed object, but is essentially independent of the object's visual features, including its size, shape, color and texture (Flombaum, Scholl & Santos, 2008). Object files are, indeed, instantiated for moving patterns that are indistinguishable from the local background when stationary (Gao & Scholl, 2010). Experiments with point-light walker displays demonstrate that human subjects can group multiple trajectories of essentially featureless objects into a single complex trajectory that is identifiable not only as that of a human being but as that of a human being performing some specific action (reviewed by Blake & Shiffrar, 2007). Experiments with animations of simple geometric shapes similarly indicate that human subjects can infer intentions, beliefs and primary emotions such as fear from the relationships between the trajectories of essentially featureless objects (reviewed by Scholl & Tremoulet, 2000; Scholl & Gao, 2013). These and other data are consistent with a model of the object file of a complex, multi-component object as a transient activation of a hierarchy of specific visuo-motor trajectory representations (Fields, 2011a). Binding of feature information such as size, shape, color and texture as well as interpretative or categorical information such as identified actions or intentions stabilizes this transient activation in visual short-term memory. A key component of this "specific trajectory recognition" model is the subsequent suppression of trajectory information in favor of feature information both in the instantiation of individual-specific object tokens representing identified objects as participants in remembered events and in the consolidation of long-term episodic memories encoding such events (Fields, 2011a; Fields, 2012b). The hierarchical nature of object files for complex objects with bound features has recently been directly demonstrated by Valdés-Sosa, Iglesias-Fuster and Torres (2014), who used a modified Navon task to show that object files representing components of a composite object can interfere with the object file representing the object as a whole.

If the most fundamental representation of objecthood for human beings is a hierarchy of specific trajectories, one can ask whether these hierarchies are learned or innate. The striking demonstration by Simion, Regolin and Bulf (2008) that newborns could distinguish upright from inverted point-light walkers suggests that at least some specific trajectory recognition networks are active at birth (but see Pavlova, 2012 for a review of contrasting findings). Early activity of trajectory recognition networks is also indicated by the higher salience of motion compared to static features such as size, shape, color and texture – with the prominent exception of the features defining human-like faces – for infants from the earliest ages tested (reviewed by Gerhardstein *et al*., 2009). However, prior to about two years of age infants appear to employ different criteria for inferring objecthood from trajectory information than do older children or adults (Bremner *et al*., 2005; 2007); hence the specificities of at least some trajectory recognition networks change during early development, presumably as a consequence of continuing visual experience with moving objects. However, ethical considerations prevent the kinds of manipulations of the environments of human infants that might reveal how visual experience enables or refines the specificities of trajectory-recognition networks or the criteria employed for judgments of object persistence. We do not know, for example, whether infants reared in environments lacking specific kinds of motion stimuli – for example, infants who never saw objects briefly disappear behind occluders – would develop specific dysfunctions in object identification.

Developmental robotics platforms such as the iCub robot and the associated robot-environment virtual-reality simulation system (Sandini, Metta & Vernon, 2007; Vernon, von Hofsten and Fadiga, 2011) allow the replication in a fully-controlled system of classic experiments characterizing infant object identification (e.g. Chen & Weng, 2004; Schlesinger, Amso, Johnson, Hantehzadeh & Gupta, 2012). They also allow the separate investigation of two kinds of interactions with objects that cannot ethically be drastically manipulated in human infants: exploratory manipulation and joint attention. Infants learn by grasping, pulling and pushing that some visible parts of the world, e.g. their own limbs, items of clothing or small toys, are separately movable while others, e.g. large items of furniture or walls of rooms, are not. The separately movable things acquire visual trajectories as a result of such manipulations and are represented and remembered as individual manipulable objects or as independently-manipulable components of larger individual objects (reviewed by von Hofsten, 2007). This manipulative object-learning strategy has been implemented on the iCub by Nguyen *et al.* (2013), who showed that curiosity-driven manipulation significantly improves learning performance over static displays or even manipulations by teachers. Infants also learn by employing a rapidly-developing capability for gaze following and joint attention (reviewed by Tomasello *et al*., 2005) to track object manipulations performed by others while segregating the "teacher" from the object being manipulated. Experiments with the iCub have shown both that object manipulation by a teacher facilitates object learning and that the objects being manipulated can be correctly segregated from the manipulating teacher (e.g. the teacher's hand), especially if they are independently manipulable by the robot (Lyubova & Filliat, 2012; Nguyen *et al.*, 2013).

The objects manipulated by infants or robots are inevitably featured: they have particular sizes, shapes, textures, colors, etc. The salience of such features to human infants develops gradually over the first two years, as does the ability to recognize physical constraints pertaining to such features, such as the inability of a large object to fit inside a small container (reviewed by Baillargeon, 2008; Baillargeon *et al.*, 2012). Infants are, however, visually sensitive from birth to features in general, just as they are to the complex feature aggregates defining faces, and they are capable of feature-based object categorization by as early as four months (reviewed by Rakison and Yermoleva, 2010). Sensitivity, therefore, does not by itself imply salience. The salience of non-facial features for infants can be easily manipulated in the laboratory; for example, Needham, Dueker and Lockhead (2005) showed that four-and-a-half month old infants do not segregate unfamiliar, unmoving objects by shape, but are capable of doing so after being shown similar objects moving. The developmental lag between sensitivity to object features and their salience, together with the modifiability of feature salience by perceived motion, suggests that feature binding to the object file, not just feature detection and processing *per se*, develops in stages over the first two years. Typical differences between infant and adult categories (reviewed by Mandler, 2004) may, moreover, at least in part reflect differences in trajectory-based object-file instantiation criteria (Bremner *et al*., 2005; 2007) and hence differences in what "counts as" an object.

Category learning dramatically accelerates with the onset of language learning; by the preschool years, categories can be taught by exemplification followed by verbal definition as they are to adults. It is not until this point or even later that the implicit, motion-based criteria of objecthood are supplemented by a consciously-accessible verbal concept of objecthood (reviewed by Fields, 2013b) or that the general, abstract notion of an object becomes available. The facilitation of category learning by language has been demonstrated using both game-theory methods and developmental robotics (reviewed by Cangelosi, 2010). Such studies have, so far, generally made the assumption that all players in the simulation see the same objects and see them in the same way up to point of view. Narens, Jameson, Komarova and Tauber (2012) relax this assumption somewhat in game-theoretic studies of shared-language development in a one-dimensional color world by allowing agents to group colors into named categories differently; all agents in these simulations are, however, forced by the design to employ a shared categorization dimension. Studies that relaxed the assumption of overlapping classifications still further by decoupling feature binding from feature detection and explicitly modeling the relationship between binding and salience could provide new insights into the mechanisms required to get category learning – and hence language learning – off the ground.

While infant and early-childhood abilities to identify and categorize objects during brief perceptual encounters have been intensively studied, the ability to re-identify particular objects as persistent individuals across substantial gaps in observation has received far less attention. Individual re-identification, as opposed to mere categorization by type, is clearly a critical skill for human beings and one that is developed from early infancy; even very young infants can distinguish their mothers from strangers, and by two years of age most infants are capable of individually re-identifying multiple family members, pets, toys, items of furniture and other objects. The general assumption that this ability to re-identify individuals is enabled by and dependent upon feature stability within individuals and feature differences between individuals has not been rigorously tested by experiments on either human infants – where ethical considerations would block many informative manipulations – or robots. Three experiments in particular, however, demonstrate that by about four year of age, children are able to use information about the causal histories of identically-featured objects to distinguish between them, even if the relevant histories have not been fully observed, (Gutheil et al., 2008; Hood & Bloom, 2008; Frazier & Gelman, 2009). Such experiments, as well as common observations that children, like adults, can often not only distinguish between identically-featured individuals, but also re-identify individuals following unobserved histories involving substantial changes in their features, forcefully raise the questions of how causal reasoning is employed to make such distinctions and what kinds of causes are typically employed in such reasoning.

Behavioral and neurocognitive data from adults suggest a model in which structure-mapping inferences (Gentner, 1983) implemented by the pre-motor action planning system are critical to the use of causal constraints in judgments of individual identity across extended gaps in observation (Fields, 2012a; 2013a). Employing this model in a developmental context, however, requires incorporating both default solutions and the constraints imposed by age-limited causal-reasoning capabilities. The human bias toward agentive causation, first demonstrated in the laboratory by Heider and Simmel (1944) and since studied extensively by many investigators (reviewed by Scholl & Tremoulet, 2000; Scholl & Gao, 2013) provides a case in point. Human infants perform goal-directed actions and appear to experience a sense of self as agent from birth (reviewed by Rochat, 2012). Infants are capable of recognizing agency and inferring goals from perceived actions (reviewed by Luo & Baillargeon, 2010; Csibra & Gergely, 2012) and interpret causation as due to agency by default. While neither the implementation nor the development of these capabilities are fully understood, substantial evidence supports the early development of visuo-motor mirror networks that bidirectionally associate representations of observed actions with action plans and hence with typical action goals (reviewed by Rizzolatti and Craighero, 2004; Craighero, Metta, Sandini and Fadiga, 2007; Rizzolatti and Sinigaglia, 2010), and it is reasonable to assume that such representations engage a felt sense of agency. By about four years old, at least some children are able to attribute independent causality to mechanisms as opposed to agents (Sobel *et al*., 2007; Sobel & Buchanan, 2009), an ability that may reflect experience-dependent alterations in the specificity of mirror-network components (reviewed by Heyes, 2010). However, a substantial bias toward agency remains in the causal reasoning of most children well into adolescence (e.g. Kelemen, 2004; Bloom, 2007) and characterizes the causal reasoning even of adult physical scientists when they are time-stressed (Kelemen, Rottman & Seston, 2013). One can hypothesize, therefore, that before four years of age, any causal models used to infer individual identity across feature changes or in the presence of identically-featured competitors are based exclusively on causation by agents, and that even in older children, such identity-preservation models remain biased toward causation by agency. If this is the case, the human representation of individual identity over time is strongly developmentally coupled to the representation of agency, and may be expected to remain so throughout the lifespan (Fields, 2014).

While intention inference and intention sharing have been demonstrated in robotic models (e.g. Dominey & Warneken, 2011; Lallée *et al.*, 2012), the demonstrated tasks did not require a discrimination between "agents" as entities capable of acting intentionally and "non-agents" incapable of acting intentionally and hence did not directly address the representation of agency. The development of robotic or other biologically-realistic computational models capable of distinguishing agents from non-agents based on their features or motion but also capable of exhibiting human-like performance on a Heider-Simmel agency-attribution task or of exhibiting human-like performance degradation in the "wolfpack" task of Gao, McCarthy and Scholl (2010) could provide substantial insights into the human construction of agency as a characteristic of some but not all objects and a cause of some but not all events. Such systems would provide, in turn, a basis for exploring both the development of causal models based on attributions of agency and the use of such models to re-identify specific individual objects over extended time periods during which their locations or features were altered.

In summary, the inference of object identity over short, within-scene times by both humans and robots is increasingly well understood. Available robotic or, more generally, biologically-inspired cognitive architectures (BICAs) do not, however, yet fully replicate the human object-file instantiation followed by feature binding process (reviewed by Goertzel, Lian, Arel, de Garis & Chen, 2010), and hence can shed only limited light on the developmental progression, normal-range variation or failure modes of this process. The LIDA architecture (Franklin *et al.*, 2007), for example, includes separate models of the dorsal and ventral visual streams, but the model of the dorsal stream only drives reflexive action generation and does not contribute to object identification (Snaider, McCall & Franklin, 2011; Franklin, Madl, D'Mello & Snaider, 2014). The MACSi architecture employed in the experiments of Nguyen *et al.* (2013) goes considerably farther toward modeling the human visual processing pathway, using motion cues for initial object segmentation and shape determination with depth and color feature binding performed subsequently (Ivaldi *et al*., 2012; Lyubova & Filliat, 2012). The perceptual-completion model of Schlesinger, Amso and Johnson (2007), used in the iCub experiments of Schlesinger *et al*. (2012), goes farther still, implementing intensity, motion, color and orientation detection as parallel pathways with subsequent binding to construct a "salience map" of the input (see also Schlesinger, 2013 for an extended discussion of this model). While artificial visual systems have achieved considerable robustness, however, key questions about individual object re-identification, including the basic question of the minimal required *a priori* information, remain to be answered.

Extending current developmental robotic platforms or other BICAs toward more detailed models of the human visual, visuomotor and visual categorization systems would enable experiments that cannot be performed on human infants to be performed on robot surrogates. As an example, a system that explicitly modeled the formation of categorized object tokens from fully-bound object files could be used to systematically examine the dependence of categorization and hence object-token structure on the relative strengths of dorsal-stream motion inputs versus ventral-stream feature inputs. Imbalancing this system in favor of dorsal-stream inputs during early infancy could be expected to produce motion-based object categories that would cross-cut typical human feature-based categories and disallow feature-based identification of objects that moved differently during different observed episodes. A system deploying such motion-based categories could be expected to exhibit many of the cognitive and behavioral symptoms associated with autism-spectrum disorders, including social interaction and language learning deficits (Fields, 2012b); whether this is the case cannot, however, be tested directly with human infants.

The processes by which individual objects are re-identified across substantial gaps in observation, particularly in the presence of significant feature change or similarly-featured competitors, are moreover poorly understood even in adult humans. It is clear that individual object re-identification across substantial gaps in observation at least sometimes requires the use of causal reasoning to construct a plausible "story" about how the object could have come to be where it is and have come to exhibit the altered or different features that it exhibits. Such causal reasoning often, but not always, involves either the assumption that the object in question is itself an agent, or the postulation of one or more agents that have moved or altered the object. Robotic systems implementing experimentally-manipulable causal reasoning systems may provide useful testbeds for studying the role of causal reasoning in individual re-identification over extended time.

Epistemic feelings as intrinsic motivators for individual object re-identification

In everyday human experience, most objects are categorized by type and most previously-encountered individuals are re-identified as such rapidly and unconsciously. It does not, for example, typically require an extended period of conscious deliberation to re-identify one's spouse, house, car or colleagues. In some cases, however, categorization or individual re-identification occur slowly and do involve conscious deliberation. As Eichenbaum, Yonelinas and Ranganath (2007) emphasize in their review of the cognitive neuroscience of recognition, when these processes are sufficiently temporally extended to involve consciousness, they tend to be accompanied by distinctive feelings including familiarity, puzzlement, curiosity and finally, when categorization or re-identification is achieved, confidence or even certainty. While the production of these "epistemic" (Arango-Muñoz, 2014) feelings is not well understood, they appear to characterize three distinct phenomenal stages of the recognition or re-identification process: the feeling of familiarity appears to enable recognition or re-identification by focusing attention on a particular object as potentially recognizable or re-identifiable, puzzlement and curiosity appear to motivate the search for relevant episodic memories or general knowledge, and confidence in or certainty about a categorization or re-identification solution appear to terminate the process and foreclose searches for alternative solutions. The generality of this sequence of feelings across problem solving domains suggests that when recognition or re-identification are sufficiently temporally extended to involve consciousness, they share many of the characteristics of others kinds of conscious, deliberate problem solving.

From a formal, mathematical perspective, uncertainty is a measure of the mismatch between expected and actual states of the world. In this formal sense, uncertainty minimization can be viewed as a criterion of successful behavior; indeed Friston (2010) has shown that under reasonable information-theoretic assumptions, the maintenance of homeostasis corresponds to the minimization of uncertainty. The relationship between uncertainty in this formal sense and uncertainty as experienced is clearly complex, as the common occurrence of subjects being fully certain of beliefs that are demonstrably false demonstrates. Given the increasingly well-recognized functional integration between affective and cognitive systems (reviewed by Pessoa, 2008; Inzlicht, Bartholow and Hirsh, 2015), it is reasonable, however, to assume that experienced uncertainty generally reflects an actual lack of relevant knowledge and serves as a motivator of consciously-directed knowledge-seeking behaviors including problem solving. Experienced uncertainty or risk can be accompanied by feelings of discomfort, anxiety or fear and motivate avoidance, but can also be accompanied by more positive feelings ranging from vague restlessness to interest or attractiveness and motivate approach and exploration. Responsiveness to these latter motivations constitutes curiosity (reviewed by Kashdan &

Silvia, 2009; Silvia, 2012). The affective correlates of curiosity are complex and even paradoxical; curiosity may be accompanied by feelings of either excitement or apprehension, and even by a distinct sense of risk or danger that is nevertheless compellingly attractive. Controversy remains as to whether curiosity is better conceptualized as fundamentally aversive – as something to be relieved – or as fundamentally positive. Litman and Jimerson (2004) proposed that common usage conflates two distinct forms of curiosity, an aversive form reflecting a task-specific need for information and a positive form reflecting non-task-specific interest (see also Litman, 2010); this distinction accords well with the distinction between extrinsic and intrinsic motivations for information seeking (e.g. Gottlieb, Oudeyer, Lopes & Baranes, 2013; Cangelosi & Schlesinger, 2015). Human curiosity appears to be just one instance of a pan-mammalian and perhaps pan-vertebrate emotional response that motivates exploratory behavior and pursuit of rewards (Panksepp, 2005). One might speculate that the sense of attractive uncertainty expressed as curiosity may, together with the contrasting sense of adversive uncertainty or risk, represent the evolutionary roots of epistemic feelings in general.

In addition to their role as motivators of problem solving, epistemic feelings play a critical and often-overlooked role in regulating and terminating problem solving. Feelings of confidence and puzzlement, for example, are plausibly linked to the absence or presence, respectively, of representational or interprocess conflict detected by the anterior cingulate and communicated to the reward system (Kaplan and Oudeyer, 2007). Feelings of knowledge or certainty serve as markers that identify solutions, thus terminating problem-solving. This is particularly clear in recombinational models of creative problem solving, in which the “answer” is a novel and by definition unexpected conceptual combination that must be identified and marked as a solution in order for the recombination process to halt (Fields, 2004).

Positive affect associated with epistemic feelings indicating successful problem solving – as experienced in *Aha!* moments – may also motivate further problem seeking and solving behaviors via positive feedback to the reward system (Fields, 2011b). Problems to be solved are instances of at least short-term uncertainty, even if they are also potentially means of reducing longer-term uncertainty; hence problem *seeking* behavior can be viewed as behavior in pursuit of, as opposed to behavior avoiding, short-term uncertainty. Some people actively seek this kind of short-term uncertainty; others do not. Higher uncertainty tolerance is a defining characteristic of “openness” as a personality measure (McCrae & John, 1992) and is characteristic of creativity across disciplines (e.g. Feist, 1998). While the organic bases for differences in uncertainty tolerance and hence uncertainty-seeking behaviors across individuals remain poorly understood, a variety of studies demonstrate significant structural and functional differences in the neurocognitive systems involved in emotional responses and current-state to goal-state conflict resolution that correlate with uncertainty tolerance (reviewed by Jost & Amodio, 2012).

As a variety of conscious, deliberate problem-solving, the process of individual object re-identification can generate particularly strong epistemic feelings. From an ecological perspective, this is not surprising; different individuals within a single category – different people, for example, or different cars or houses – may offer very different affordances to the perceiver. In humans, feelings of familiarity with people or other objects are generated during ventral-stream visual processing prior to full object categorization or individual re-identification (reviewed by Eichenbaum, Yonelinas and Ranganath, 2007; Zimmer & Ecker, 2010). Familiarity with an object generally implies familiarity with its affordances; hence perceiving a familiar object can be expected to activate representations of its affordances across both cortical (reviewed by Thill *et al*., 2013) and cortical-subcortical (reviewed by Caligiori *et al*., 2013) networks, the latter involving affective systems that may contribute to the generation of feelings of familiarity. Feelings of familiarity can induce curiosity and motivate exploration, as when one encounters a familiar but not-yet-identifiable person (e.g. Eichenbaum, Yonelinas and Ranganath, 2007). Even more striking are the immediate and intense feelings of unfamiliarity and anxiety that can be generated when an unrecognized person behaves with complete familiarity, thus generating enormous and generally aversive uncertainty about the unknown person's intentions. Capgras syndrome, in which close family members are regarded as imposters, provides a compelling example, one traceable to dysfunctional emotional responses to familiar faces (reviewed by Coltheart, Langdon and McKay, 2011). The typically-deep feeling of certainty about self-identity – implicitly, about self-identity through time – forms the basis for anti-skeptical arguments by philosophers from Descartes to G. E. Moore. While temporary disruptions of the sense of self-identity can be associated with strong positive affect in meditative (e.g. Josipovic, 2014) or entheogen-induced (e.g. Carhart-Harris *et al*., 2012) experiences, chronic disruptions of the sense of self-identity in post-traumatic stress disorder (PTSD; Wolf *et al.*, 2012), other dissociative disorders (Spiegel *et al*., 2011) or Alzheimer's disease (Jicha & Carr, 2010) typically induce fear, anger, depression or other states involving strong negative affect. Loss of the ability to re-identify previously-familiar objects in Alzheimer's disease and other memory disorders similarly induces negative affect and depressed or erratic behavior. As noted earlier, dysfunctions in feature-based object re-identification, including re-identification of other people, early in development may manifest clinically as autism spectrum disorders, with their typical disruptions of affect (Fields, 2012b).

From an implementation perspective, the relevant questions are first, how epistemic feelings are generated and how they affect both the initiation and the termination of human problem-solving, including human individual object-identity determination, and second, how both their generation and functional role in problem solving and object-identity determination can be replicated in artificial systems, specifically autonomous robots that require an ability to detect, manage and even exploit environmental uncertainty. The generation of epistemic feelings associated with memories or judgments is commonly modeled as a metacognitive process that maximizes some measure of coherence across a large sample of relevant beliefs (e.g. Koriat, 2012). Both the close association of epistemic feelings with the primary emotions of fear and pleasure and the general integration of affect with cognition, however, suggests that epistemic feelings directly involve affective and reward processes as well. Considerable evidence now indicates that in humans, insular cortex re-represents inputs from the autonomic nervous system to the brainstem as affective bodily feelings including pain, hunger, nausea and many others (reviewed by Craig, 2002; 2009). The insula are also involved in integrating these interoceptive feelings with other sensory and emotional inputs (reviewed by Craig, 2010; Gu, Hof, Friston & Fan, 2013). Novel and unexpected epistemic feelings generated by insular-cortex activity can dominate longstanding metacognitive beliefs. Picard (2013), for example, reports the transient occurrence of feelings of unquestionable certainty in the existence of God as an omnipotent designer of the universe, accompanied by intense positive affect, in a self-described atheist during a right-insular cortex seizure. Using blurred images as an uncertainty-inducing stimulus and fMRI to detect task-specific activation, Jepma, Verdonschot, van Steenbergen, Rombauts and Nieuwenhuis (2012) were able to directly demonstrate activation of both right anterior insular and bilateral anterior cingulate cortex in subjects who reported experiencing curiosity. Although Jepma *et al*. (2012) interpreted the insular activation solely in terms of arousal, it is interesting to note that anterior insular cortex is broadly implicated in the integration of epistemic feelings and *right* anterior insular cortex is broadly implicated in the integration of aversive feelings in particular (e.g. Craig, 2010), consistent with the use of a blurred-image paradigm in these experiments.

While metacognitive processing has been extensively studied by AI researchers for several decades, e.g. in the context of truth maintenance systems (Doyle, 1979; de Kleer, 1986), the potential regulatory role of affect is only beginning to attract attention. The use of internally-generated measures of novelty or familiarity, predictive success or failure, and uncertainty reduction or learning rate as effectively sensory inputs that motivate either exploration or avoidance of places, situations, actions or objects has been studied primarily in a developmental robotics context (reviewed by Oudeyer & Kaplan, 2007; Gottlieb, Oudeyer, Lopes & Baranes, 2013; Cangelosi & Schlesinger, 2015). The measures of learning rate or of change in prediction uncertainty discussed by Oudeyer, Baranes and Kaplan (2013) as models of curiosity, for example, may be considered to be formally metacognitive, but their real-time functions as both motivators of exploratory behavior and directors of focused attention are analogous to and partially model the real-time functions of affectively-linked regulatory inputs – in this case, current-state to goal-state disparity inputs from the anterior cingulate cortex (e.g. Holroyd & Coles, 2008; Quilodran, Rothe & Procyk, 2008) – in humans. Outside of the relatively protected environment of infancy and early childhood, however, uncertainty measures are often also measures of physiological or social risk. Environmental indicators of such risks may be very subtle; indeed a primary idea motivating the "social brain" hypothesis is that evolutionary arms races between conspecifics will lead inevitably both to increasingly hidden indicators of risk and increasingly greater cognitive resources dedicated to risk detection (reviewed by Dunbar, 2003; Dunbar & Shultz, 2007). The ways in which affective indicators of risk, and in particular interoceptive indicators of risk, influence human cognition remain largely uncharacterized, as does the range of environmental conditions to which such signals are responsive. As a primary function of interoceptive signals is the maintenance of bodily homeostasis (Craig, 2002; 2009), one could hypothesize that *any* environmental condition that tended to reproducibly affect bodily homeostasis in the ancestral human niche might be detected by some bodily system and represented by some interoceptive signal. It is straightforward to represent current power usage or remaining battery life in the operating system of a robot, but what other homeostasis-relevant signals might autonomous robots embedded in physically or socially risky environments usefully represent? Understanding how such cues can be detected, processed and represented in a way that productively influences behavior may, however, prove to be important for understanding human uncertainty detection and management, both in decision making about object identity and decision making in general.

As in the case of object tracking through time discussed above, robotic platforms may provide opportunities for investigating the roles of epistemic feelings as both motivators and terminators of individual identification processes in ways that would be infeasible or unethical if applied to human subjects. Reward, for example, is standardly treated as a single number that can vary up or down in consequence of a robot's actions. In humans, however, situations can be attractive and aversive at the same time, and actions can have positive consequences along some dimensions but negative consequences along others. Different reward dimensions may also interact in complex ways. For example, the consequences of mis-identifying or incorrectly judging the familiarity of a correctly categorized individual may be more negative in some cases than the consequences of mis-categorization, but less negative in other cases. A multidimensional, vector model of reward for robot tasks would allow essentially arbitrary manipulation of the reward landscape as a function of the context in which an action is taken and hence allow more detailed studies of, for example, contexts in which risks are high but potential benefits are high as well. Such reward landscapes have been studied in the abstract (e.g. Brázdil, Jančar & Kučera, 2010; Chatterjee, Randour & Raskin, 2014), but their potential for application to autonomous robots that must explore and learn from high-risk task environments remains to be investigated.

A second, related area of opportunity is the explicit modeling of the inputs from the body that, in humans and presumably in other animals, are re-represented as affective experiences. For animals, the maintenance of homeostasis in the face of changing task demands requires the reallocation of energy resources to processes including motion, sensation (i.e. information acquisition), cognition, metabolism and repair. Such processes produce feedback signals critical for regulation. In robots, not just total power usage or remaining battery life, but also specific resource allocations to individual computational and motor processes could be monitored and employed as surrogates for such signals (for some initial exploration along these lines, see Kernbach and Kernbach, 2011). Manipulations of a multidimensional reward landscape could then model variations in the perceived difficulty of task components in the absence of significant variations in actual computational or mechanical complexity.

Finally, robots or even BICAs without explicitly modeled "bodies" may provide useful and fully-manipulable models of cognitive biases. The sudden replacement of anxiety by certainty in delusional patients, for example, is often conceptualized in terms of a jumping to conclusions bias (e.g. Freeman, 2007), a condition also linked to non-clinical deficiencies in causal reasoning (Brosnan, Ashwin & Gamble, 2013). Cognitive biases are, however, useful and often essential in high-uncertainty, resource-limited task environments. Individual object re-identification uncertainty, in particular, cannot *objectively* be reduced to zero; as noted earlier, all individual re-identification is heuristic. As Oudeyer, Baranes and Kaplan (2013) point out, autonomous systems capable of free exploration run the risk of devoting too many attentional resources to the resolution of uncertainties that cannot feasibly be resolved and do not, in the current context, need to be. The trick, of course, is in knowing when they need to be. In humans, subtle epistemic feelings often provide warnings that what appear to be innocuous "sleeping dogs" are not sleeping and may not be dogs; such sensitive detectors of mis-identification risk may be needed by autonomous robots as well.

Conclusion

What is required to assess the opportunities and risks afforded by a task environment sufficiently quickly and accurately to guide productive behavior? An ability to fit objects or events into categories with which opportunities and risks have already been associated is clearly advantageous. As the number of relevant categories and the subtlety of the distinctions between them become large, however, the categorization process becomes increasingly resource intensive. An ability to identify particular individuals with which extensive categorical knowledge has already been associated shortcuts this resource-intensive process. One would, therefore, expect organisms inhabiting complex environments to devote cognitive resources to both the identification and characterization of specific individuals and their re-identification over time, and one would expect such abilities to benefit artificial systems operating in similar environments.

Gaps in observation pose a substantial problem for any system attempting to re-identify individual objects over time. Human beings display a remarkable ability to surmount this problem, regularly re-identifying objects that they have not observed for hours, days or even many years. They often do this, moreover, in the face of significant changes in the features of the re-identified individuals, and in the presence of multiple similarly-featured competitors. This ability is widely taken for granted, and has received relatively little direct attention from researchers. The development of mobile autonomous robots for tasks that require long-term social interaction either with humans or among themselves brings the need to understand the functional underpinnings of individual object re-identification to the fore. As indicated in the present review, much remains to be done.

*****
Box 2 about here
*****

Uncertainty about affordances is an inevitable consequence of uncertainty about individual object re-identification. Here it is clear that learning must play a dominant role: infants may be born able to recognize certain categories, but they must learn the reliably identifying features, behaviors, and other characteristics of every individual in their environments that is recognizable as such. How they do this remains largely unknown. By permitting essentially arbitrary manipulation of “prior knowledge,” developmental robotics provides an opportunity to formulate and test models that extend well beyond the ethical or even feasible bounds of developmental psychology. Such models may provide new insights not just into the mechanisms at work in human development, but also into the evolutionary and cultural reasons why humans display such extraordinary diversity in uncertainty tolerance, curiosity, and problem-seeking exploratory behavior.

Conflict of Interest Statement

The author states that he has no conflicts of interest relevant to the reported research.

References

Arango-Muñoz, S. (2014). The nature of epistemic feelings. *Philosophical Psychology*, 27, 193-211.

Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y., Ogino, M. & Yoshida, C. (2009). Cognitive developmental robotics: A survey. *IEEE Transactions in Autonomous Mental Development* 1, 12-34.

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

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.

Blake, R. & Shiffrar, M. (2007). Perception of human motion. *Annual Review of Psychology*, 58, 47-73.

Bloom, P. (2007). Religion is natural. *Developmental Science*, 10, 147-151.

Brázdil, T., Jančar, P. & Kučera, A. (2010). Reachability games on extended vector addition systems with states. *Automata, Languages and Programming: Lecture Notes in Computer Science*, 6199, 478-

489.

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.

Brosnan, M., Ashwin, C. & Gamble, T. (2013). Greater empathizing and reduced systemizing in people who show a jumping to conclusions bias in the general population: Implications for psychosis. *Psychosis*, 5, 71-81.

Caligiori, D., Pezzulo, G., Miall, R. C. and Baldassarre, G. (2013). The contribution of brain sub-cortical loops in the expression and acquisition of action understanding abilities. *Neuroscience and Biobehavioral Reviews*, 37, 2504-2515.

Cangelosi, A. (2010). Grounding language in action and perception: From cognitive agents to humanoid robots. *Physics of Life Reviews*, 7, 139-151.

Cangelosi, A. & Schlesinger, M. (2015). *Developmental Robotics: From Babies to Robots*. Cambridge, MA: MIT Press.

Carhart-Harris , R. L., Erritzoe, D., Williams, T., *et al*. (2012). Neural correlates of the psychedelic state as determined by fMRI studies with psilocybin. *Proceedings of the National Academy of Sciences U.S.A.*, 109, 2138-2143.

Chatterjee, K., Randour, M. & Raskin, J.-F. (2014). Strategy synthesis for multi-dimensional quantitative objectives. *Acta Informatica*, 51, 129-163.

Chen, Y. & Weng, J. (2004). Developmental learning: A case study in understanding "object permanence." In L. Berthouze, H. Kozima, C. G. Prince, *et al*. (Eds) *Proceedings of the Fourth International Workshop on Epigenetic Robotics* (pp. 35-42). Lund: Lund University.

Coltheart, M., Langdon, R. and McKay, R. (2011). Delusional belief. *Annual Review of Psychology,* 62, 271-298.

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Reviews Neuroscience,* 3, 655-666.

Craig, A. D. (2009). How do you feel – now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, 10, 59-70.

Craig, A. D. (2010). The sentient self. *Brain Structure and Function*, 214, 563-577.

Craighero, L., Metta, G., Sandini, G. and Fadiga, L. (2007). The mirror-neurons system: Data and models. *Progress in Brain Research*, 164, 39-59.

Csibra, G. & Gergely, G. (2012). Teleological understanding of actions. In M. R. Banaji & S. A. Gelman (Eds) *Navigating the Social World: What Infants, Children, and Other Species Can Teach Us* (pp. 38-43). Oxford: Oxford University Press.

de Kleer, J. (1986). An assumption-based TMS. *Artificial Intelligence*, 28, 127-162.

Dietrich, E. & Fields, C. (1996). The role of the Frame Problem in Fodor's Modularity thesis: A case study in rationalist cognitive science. In K. M. Ford and Z. W. Pylyshyn (Eds) *The Robot's Dilemma Revisited: The Frame Problem in Artificial Intelligence* (pp. 9-24). Norwood, NJ: Ablex.

Dominey, P. F. & Warneken, F. (2011). The basis of shared intentions in human and robot cognition. *New Ideas in Psychology*, 29, 260-274.

Doyle, J. (1979). A truth maintenance system. *Artificial Intelligence,* 12, 231-272.

Dunbar, R. I. M. (2003). The social brain: Mind, language and society in evolutionary perspective. *Annual Review of Anthropology*, 32, 163-181.

Dunbar, R. I. M. & Shultz, S. (2007). Evolution in the social brain. *Science*, 317, 1344-1347.

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

Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review*, 2, 290-309.

Fields, C. (2004). The role of aesthetics in problem solving: some observations and a manifesto. *Journal of Experimental and Theoretical Artificial Intelligence*, 16, 41-55.

Fields, C. (2011a). 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. (2011b). From "Oh, OK" to "Ah, yes" to "Aha!": Hyper-systemizing and the rewards of insight. *Personality and Individual Differences*, 50, 1159-1167.

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. (2013a). How humans solve the frame problem. *Journal of Experimental and Theoretical Artificial Intelligence*, 25, 441-456.

Fields, C. (2013b). The principle of persistence, Leibniz's law, and the computational task of object re-identification. *Human Development*, 56, 147-166.

Fields, C. (2014). Motion, identity and the bias toward agency. *Frontiers in Human Neuroscience*, 8, 597 (doi: 10.3389/fnhum.2014.00597).

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* (pp. 135-164). New York: Oxford University Press.

Franklin, S., Ramamurthy, U., D'Mello, S. K., McCauley, L., Negatu, A., Silva L., R. and Datla, V. (2007). LIDA: A computational model of global workspace theory and developmental learning. *AAAI Fall Symposium on AI and Consciousness: Theoretical Foundations.* Palo Alto: AAAI (pp. 61-66).

Franklin, S., Madl, T., D'Mello, S. & Snaider, J. (2014). LIDA: A systems-level architecture for cognition, emotion and learning. *IEEE Transactions on Autonomous Mental Development*, 6, 19-41.

Frazier, B. N. & Gelman, S. A. (2009). Developmental changes in judgments of authentic objects. *Cognitive Development*, 24, 284- 292.

Freeman, D. (2007). Suspicious minds: The psychology of persecutory delusions. *Clinical Psychology Review*, 27, 425-457.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11, 127-138.

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.

Gao, T., McCarthy, G. & Scholl, B. (2010). The Wolfpack Effect: Perception of animacy irresistibly influences interactive behavior. *Psychological Science*, 21, 1845-1853.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155-170.

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

Goertzel, B., Lian, R., Arel, I., de Garis, H. & Chen, S. (2010). A world survey of artificial brain projects, Part II: Biologically inspired cognitive architectures. *Neurocomputing* 74, 30-49.

Gottlieb, J., Oudeyer, P.-Y., Lopes, M. & Baranes, A. (2013). Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends in Cognitive Sciences*, 17, 585-593.

Gu, X., Hof, P. R., Friston, K. J. & Fan, J. (2013). Anterior insular cortex and emotional awareness. *Journal of Comparative Neurology,* 521, 3371-3388.

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.

Heider, F., and Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, 57, 243-259.

Heyes, C. (2010). Where do mirror neurons come from? *Neuroscience and Biobehavioral Reviews*, 34, 575-583.

Holroyd, C. B., & Coles, M. G. H. (2008). Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior. *Cortex*, 44, 548-559.

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

Inzlicht, M., Bartholow, B. D. and Hirsh, J. B. (2015). Emotional foundations of cognitive control. *Trends in Cognitive Sciences*, 19, 126-132.

Ivaldi, S., Lyubova, N., Gérardeaux-Viret, D. *et al*. (2012). Perception and human interaction for developmental learning of objects and affordances. *Proceedings of the 12th IEEE-RAS International Conference on Humanoid Robots* (Humanoids Japan), 1-7.

Jicha, G. A. & Carr, S. A. (2010). Conceptual evolution in Alzheimer's disease: Implications for understanding the clinical phenotype of progressive neurodegenerative disease. *Journal of Alzheimer's Disease*, 19, 253-272.

Jepma, M., Verdonschot, R. G., van Steenbergen, H. Rombauts, S. A. R. B. & Nieuwenhuis, S. (2012). Neural mechanisms underlying the induction and relief of perceptual curiosity. *Frontiers in Behavioral Neuroscience*, 6, 5 (doi 10.3389/fnbeh.2012.00005).

Josipovic, Z. (2014). Neural correlates of nondual awareness in meditation. *Annals of the New York Academy of Sciences*, 1307, 9-18.

Jost, J. T. & Amodio, D. M. (2012). Political ideology as motivated social cognition: Behavioral and neuroscientific evidence. *Motivation and Emotion*, 36, 55-64.

Kahneman, D. & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds) *Varieties of Attention* (pp. 29-61), New York: Academic Press.

Kaplan, F. and Oudeyer, P.-Y. (2007). In search of the neural circuits of intrinsic motivation. *Frontiers in Neuroscience*, 1, 225–236.

Kashdan, T. B. & Silvia, P. J. (2009). Curiosity and interest: The benefits of thriving on novelty and challenge. In C. R. Snyder & S. J. Lopez (Eds) *Oxford Handbook of Positive Psychology* (pp. 367-374) Oxford: Oxford University Press.

Kelemen, D. (2004). Are children "intuitive theists"? Reasoning about purpose and design in nature.

Psychological Science, 15, 295-301.

Kelemen, D., Rottman, J. & Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default. *Journal of Experimental Psychology: General*, 142, 10741083.

Kernbach, S. & Kernbach, O. (2011). Collective energy homeostasis in a large-scale microrobotic swarm. *Robotics and Autonomous Systems*, 59, 1090-1101.

Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological Review*, 119, 80-113.

Lallée, S., Pattacini, U., Lemaignan, S. *et al*. (2012). Towards a platform-independent cooperative human robot interaction system III: An architecture for learning and executing actions and shared plans. *IEEE Transactions on Autonomous Mental Development*, 4, 239-253.

Litman, J. A. (2010). Relationships between measures of I- and D-type curiosity, ambiguity tolerance, and need for closure: An initial test of the wanting-liking model of information-seeking. *Personality and Individual Differences*, 48, 397-402.

Litman, J. A. & Jimerson, T. L. (2004). The measurement of curiosity as a feeling of deprivation. *Journal of Personality Assessment*, 82, 147-157.

Luo, Y. & Baillargeon, R. (2010). Toward a mentalistic account of early psychological reasoning. *Current Directions in Psychological Science*, 19, 301-307.

Lyubova, N. & Filliat, D. (2012). Developmental approach for interactive object discovery. *Proceedings of the 2012 International Joint Conference on Neural Networks* (IJCNN), 1-7.

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

McCarthy, J. & Hayes, P. J. (1969). Some philosophical problems considered from the standpoint of artificial intelligence. In B. Meltzer and D. Ritchie (Eds) *Machine Intelligence*, Vol. 4 (pp. 463-502). New York: Elsevier.

McCrae, R. R. & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60, 175-215.

Narens, L., Jameson, K. A., Komarova, N. L. and Tauber, S. (2012). Language, categorization and convention. *Advances in Complex Systems*, 15, Article 1150022.

Needham, A., Dueker, G. and Lockhead, G. (2005). Infants’ formation and use of categories to segregate objects. *Cognition*, 94, 215–240

Nguyen, S. M., Ivaldi, S., Lyubova, N. *et al*. (2013). Learning to recognize objects through curiosity-driven manipulation with the iCub humanoid robot . *Proceedings of the 2013 IEEE Third Joint International Conference on Development and Learning and Epigenetic Robotics* (ICDL), 1-8.

Oudeyer, P.-Y. & Kaplan, F. (2007). What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics*, 1, 6 (doi 10.3389/neuro.12.006.2007).

Oudeyer, P.-Y., Baranes, A. & Kaplan, F. (2013). Intrinsically motivated learning of real world sensorimotor skills with developmental constraints. In G. Baldassarre & M. Mirolli (Eds) *Intrinsically Motivated Learning in Natural and Artificial Systems* (pp. 303-365). Berlin: Springer.

Panksepp, J. (2005). Affective consciousness: Core emotional feelings in animals and humans. *Consciousness and Cognition*, 14, 30-80.

Pavlova, M. A. (2012). Biological motion processing as a hallmark of social cognition. *Cerebral Cortex*, 22, 981-995.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, 9, 148-158.

Picard, F. (2013). State of belief, subjective certainty and bliss as a product of cortical dysfunction. *Cortex*, 49, 2494-2500.

Quilodran, R., Rothe, M., & Procyk, E. (2008). Behavioral shifts and action validation in the anterior cingulate cortex. *Neuron*, 57, 314-325.

Rakison, D.H. & Yermoleva, Y. (2010). Infant categorization. *Wiley Interdisciplinary Review: Cognitive Science*, 1, 894-905.

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

Rizzolatti, G., and Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience*, 11, 264-274.

Rochat, P. (2012). Primodial sense of embodied self-unity. In V. Slaughter & C. A. Brownell (Eds) *Early Development of Body Representations* (pp. 3-18). Cambridge, UK: Cambridge University Press.

Sandini, G., Metta, G. & Vernon, D. (2007). The iCub cognitive humanoid robot: An open-system research platform for enactive cognition. In M. Lungarella, F. Iida, J. Bongard & R. Pfeifer (Eds) *50 Years of Artificial Intelligence: Lecture Notes in Computer Science, Vol. 4850* (pp. 358-369). Berlin: Springer.

Satta, R. (2013). Appearance descriptors for person re-identification: A comprehensive review. DIEE, University of Caligari Technical Report, arxiv:1307.5748v1 [cs.CV].

Schlesinger, M. (2013). Investigating the origins of intrinsic motivation in human infants. In G. Baldassarre & M. Mirolli (Eds) *Intrinsically Motivated Learning in Natural and Artificial Systems* (pp 367-392). Berlin: Springer.

Schlesinger, M. Amso, D. & Johnson, S. P. (2007). Simulating infants' gaze patterns during the development of perceptual completion. *Proceedings of the Seventh International Conference on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, Lund University, 157-164.

Schlesinger, M. Amso, D., Johnson, S. P., Hantehzadeh, N. & Gupta, L. (2012). Using the iCub simulator to study perceptual development: A case study. *Proceedings of the 2012 IEEE International Conference on Development and Learning and Epigenetic Robotics* (ICDL), 1-6.

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

Scholl ,B. J. and Tremoulet, P. (2000). Perceptual causality and animacy. *Trends Cogn. Sci*., 4, 299–309.

Scholl, B. J. & Gao, T. (2013). Perceiving animacy and intentionality: Visual processing or higher-level judgment? In M. D. Rutherford & V. A. Kuhlmeier (Eds.) *Social Perception: Detection and Interpretation of Animacy, Agency, and Intention* (pp. 197-230). Cambridge, MA: MIT Press.

Silvia, P. J. (2012). Curiosity and motivation. In R. Ryan (Ed) *Oxford Handbook of Human Motivation* (pp. 157-166), Oxford: Oxford University Press.

Simion, F., Regolin, L. & 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.

Snaider, J., McCall, R. & Franklin, S. (2011). The LIDA framework as a general tool for AGI. In J. Schmidhuber, K. R. Thórisson & M. Looks (Eds.): *Artificial General Intelligence 2011* (*Lecture Notes in Artificial Intelligence* 6830; pp. 133-142). Berlin: Springer.

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.

Spiegel, D., Loewenstein , R. J., Lewis-Fernández , R., *et al*. (2011). Dissociative disorders in DSM-5. *Depression and Anxiety*, 28, 824-852.

Thill, S., Caligiori, D., Borghi, A. M., Ziemke, T. and Baldassarree, G. (2013). Theories and computational models of affordance and mirror systems: An integrative review. *Neuroscience and Biobehavioral Reviews*, 37, 491-521.

Tomasello, M., Carpenter, M., Call , J., Behne, T. & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28, 675-735.

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

Valdés-Sosa, M. J., Iglesias-Fuster, J. and Torres, R. (2014). Attentional selection of levels within hierarchically organized figures is mediated by object-files. *Frontiers in Integrative Neuroscience*, 8, Article 91.

Vernon, D., von Hofsten, C. and Fadiga, L. (2011). *A Roadmap for Cognitive Development in Humanoid Robots*. Berlin: Springer.

von Hofsten, C. (2007). Action in development. *Developmental Science*, 10, 54-60.

Wolf, E. J., Miller, M. W., Reardon, A. F., Ryabchenko, K. A., Castillo, D. & Freund, R. (2012). A latent class analysis of dissociation and PTSD: Evidence for a dissociative subtype. *Archives of General Psychiatry*, 69, 698-705.

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.

Text Boxes

Box 1: Glossary of terms

Categorization: The process of associating one or more descriptive categories to a perceived object on the basis of its detected features.

Epistemic feeling: A phenomenal experience indicating a presence or lack of knowledge. Examples include feelings of familiarity or unfamiliarity, certainty or uncertainty, puzzlement or curiosity.

Feature (static): A detectable characteristic such as size, shape, color or texture that an object continues to have when it is not moving relative to an observer.

Feature (motion): A detectable characteristics such as gait or speed that an object has only when moving relative to an observer.

Frame problem: The problem of determining, following an action taken in some context, which facts about the world do *not* need to be updated in consequence of the action. Following its introduction by McCarthy and Hayes (1969), the frame problem has variously been interpreted as the problem of optimally-efficient rational belief updating, the problem of optimal action-sequence planning or the problem of relevance (for a brief history, see Dietrich and Fields, 1996).

Object file: A transient working-memory representation of a time-persistent object, initially representing only location and motion, onto which static features and category membership are subsequently layered. The object file concept was introduced by Kahneman and Treisman (1984) and subsequently elaborated by Treisman and others.

Object token: A long-term memory representation of a particular, individual time-persistent object that may be referenced by multiple episodic memories involving that object.

Re-identification: The process of identifying a particular individual object as the very same individual object that was encountered at some previous time.

Box 2: Some open questions

What is the range of variation of trajectory-based criteria for object persistence during infancy? Is there significant cross-cultural variation in such criteria? Does the range of variation contract between 6 and 24 months? Do variations in trajectory-based object persistence criteria correlate with later pathology, e.g. autism or ADHD?

If features are held constant, how sensitive is object re-identification to variations in the motions executed by an object? How does such sensitivity vary with object category? How does such sensitivity vary across the developmentally-typical population and between typical development and pathology, e.g. autism or ADHD?

Do all object categories include “essential” features that objects in that category must have as necessary conditions for category membership? How do the number and kind of properties that are essential vary across categories? How do the essential features of common entry-level categories, e.g. 'house' or 'person' vary developmentally and cross-culturally? How is such variation, if it occurs, reflected in the ability to re-identify category members following featural changes?

How dependent on causal knowledge and causal reasoning ability are object identity judgments across gaps in observation? How does the tolerance for causally-inexplicable feature change vary between object categories? How does it vary by developmental age? How does it vary cross-culturally?

How robust are the individual re-identification abilities of non-human animals against location or feature changes across extended time? Can animal-human differences in object re-identification, e.g. among the higher primates, be correlated to differences in performance on other causal reasoning tasks?

How are feelings of familiarity that precede object re-identification implemented? What are the relative contributions of the number of episodic memories involving an individual object and the emotional tone and intensity of those memories to the generation of a feeling of familiarity? How is the transition from familiarity to certainty following re-identification implemented?

Do subjects that exhibit a jumping to conclusions bias or other cognitive bias in other domains exhibit the same bias when re-identifying objects?

How is the re-identification process modified in cases in which re-identification errors have serious negative consequences?

What level of agreement between robot and human performance on standard developmental psychology experiments, together with what other criteria concerning the cognitive model implemented by the robot, would enable robots to be used as surrogates for humans in experiments not performable on humans?