

EPISODIC MEMORY, PHENOMENAL EMPATHY, AND THE REPRESENTATION  
OF TIME

A thesis submitted to the faculty of  
San Francisco State University

In partial fulfillment of  
the requirements for  
the Degree

AS

36

2018

Master of Arts

PHIL

M455

In

Philosophy

by

Patrick Douglas Mellor

San Francisco, California

January 2018

Copyright by  
Patrick Douglas Mellor  
2018

## CERTIFICATION OF APPROVAL

I certify that I have read *Episodic Memory, Phenomenal Empathy, and the Representation of Time* by Patrick Douglas Mellor, and that in my opinion this work meets the criteria for approving a thesis submitted in partial fulfillment of the requirement for the degree Master of Arts in Philosophy at San Francisco State University.



Carlos Montemayor, Ph.D.  
Associate Professor



David Landy, Ph.D.  
Associate Professor

# EPISODIC MEMORY, PHENOMENAL EMPATHY, AND THE REPRESENTATION OF TIME

Patrick Douglas Mellor  
San Francisco, California  
2018

This paper will synthesize the existing literature in the fields of animal behavior and human development, episodic memory, theory of mind, and the representation of time, establishing a close relationship between the three, and giving empirically supported conceptual distinctions between two forms of episodic memory, and two forms of theory of mind. The minimal requirements for the development of each form of both episodic memory and theory of mind will be determined. Delineating the conditions for the development of each form of both episodic memory and theory of mind will show that there is a dependency relationship necessitating a sequential and alternating development of the two forms of each faculty, specifically that prospective episodic memory can develop only following the advent of both retrospective episodic memory, and mind-reading, in that order, and that metacognition can only arise subsequent to prospective episodic memory. The paper will conclude with an analysis of the relationship between episodic memory, metacognition, and temporal representation, showing that temporal representation is fundamentally conditioned by the structure of episodic memory and theory of mind.

I certify that the abstract is a correct representation of the content of this thesis.

  
Chair, Thesis Committee

12/15/2017

Date

## ACKNOWLEDGEMENTS

I would like to thank Dr. Carlos Montemayor for giving an amazing example throughout his seminars of how to approach questions in philosophy of mind with wonder, explanatory rigor, and in the spirit of universal inquiry. I will always be grateful to Dr. Abrol Fairweather for his support and encouragement throughout the last four years, without which I would not have begun this program, and to Dr. Alice Sowaal, for teaching me the true joy of the rationalist conception of time and eternity. I thank John Mellor, Tara Mellor, Victor Mercado, Claire Smith, Mark Okada, Baeo Maltinsky, and Elysha Massatt for multitudes of reasons individually and collectively. Through your loving understanding, I see myself and form my world. Finally, I thank Calliope, the first of the muses and daughter of Memory, a loving and fearless quail who, in the six years that she was with me, showed me the depth of empathy that a bird could share.

## TABLE OF CONTENTS

Introduction.....	1
Memory.....	2
1.1    Forms of Memory.....	2
1.2    The Structure of Episodic Memory.....	5
1.3    The Minimal Requirements for Episodic-like Memory in Animals.....	7
1.4    Retrospective Episodic-like Memory in Animals.....	9
1.5    Prospective Episodic-like Memory in Animals.....	11
1.6    Prospective Episodic Memory, Time, and the Agent Representation.....	13
Empathy.....	15
2.1    Prospective Episodic Memory and Theory of Mind.....	15
2.2    The Components of Theory of Mind.....	16
2.3    The Empathy-first Model.....	18
2.4    Empathy across Phylogeny.....	19
2.5    The Difficulty of Testing for Metacognition.....	20
2.6    Mind Reading and Prospective Episodic Memory.....	24
2.7    Prospective Episodic Memory and Metacognition.....	26
2.8    Prospective Episodic Memory and Metacognition in Human Evolution...	27
2.9    Metacognition and Behavioral Modernity in Humans.....	29
2.10   From Retrospective Episodic Memory to Metacognition.....	32

Time.....	35
3.1    Temporal Paradoxes, the Discrete, and the Continuous.....	35
3.2    The Evolution of Temporal Representation.....	41
3.3    Temporal Representation in Unusual Human Mental States.....	45
Conclusion.....	49
References.....	50

## **Episodic Memory, Phenomenal Empathy, and the Representation of Time**

Episodic memory and theory of mind are generally agreed to be two of the most important faculties enabling the development of the concept of the self. This conclusion is supported by strong phenomenological and empirical evidence. There has not however been sufficient work devoted to resolving the directionality of the dependencies involved.

This paper will describe and support a model of how episodic memory, theory of mind, and the self-concept can be partially dissociated from one another, concentrating on the substructure of episodic memory and theory of mind, and suggesting that each of these can be dissociated into two components, which develop in an interwoven fashion, alternating between forms of episodic memory, and of theory of mind.

The categories used will be *retrospective episodic memory* (REM), *prospective episodic memory* (PEM), *mind reading* (MR), and *metacognition* (MC), and the sequential evolutionary and ontological development of the four will be suggested to be REM → MR → PEM → MC. The self-concept will be shown to develop in two stages associated with PEM and MC respectively, and the progression from REM to MC will be shown to be accompanied by a progressive elaboration of temporal representation.

The paper will therefore also be fundamentally concerned with the evolutionary history of the development of the representation of time, and the correlate development of tensed language, and rooted in the explication of a philosophical apparatus to account for a demonstrable sequence in the evolution of these complex and generalized cognitive capacities.

## Memory

Please assume . . . that there is in our souls a block of wax . . . Let us, then, say that this is the gift of Memory, the mother of the Muses, and that whenever we wish to remember anything we see or hear or think of in our own minds, we hold this wax under the perceptions and thoughts and imprint them upon it, just as we make impressions from seal rings; and whatever is imprinted we remember and know as long as its image lasts, but whatever is rubbed out or cannot be imprinted we forget and do not know.

— Plato, *Theaetetus*

### 1.1 Forms of Memory

Memory is a fundamental requirement for most forms of behavioral flexibility, and the ability of an organism to modify their behavior in response to changes in their environment. Without some physiological system to keep track of prior states, and modify responses according to the familiarity or unfamiliarity of situations encountered, all behaviors must be hardwired, and while this does not disallow a multitude of specific responses to environmental cues, these responses will be entirely inflexible. The organism will respond the same way to a stimulus whether it is presented for the first or the thousandth time, without the possibility of learning.

Very few, if any, organisms, with the possible exception of viruses, are constructed in such a manner. While there are in any organism subsystems with a direct causal relation to external stimuli, unmediated by any updating internal process, for example reflex arcs, systems which update behaviors in response to internal modifications as a result of prior events, are almost universally present. In addition to animals, plants, protists, and prokaryotes possess such capacities. For example, *Mimosa*

*pudica*, the sensitive plant, will habituate to repeated stimuli by ceasing its characteristic leaf folding behavior, if the stimulus is a harmless touch or shake, rather than attack by a herbivorous insect (Abramson and Chicas-Mosier, 2016). The slime mold *Physarum polycephalum* will, with repeated exposure, learn that it can safely cross a bridge impregnated with a bitter substance in order to reach a food reward (Boisseau, Vogel, and Dussutour, 2016) . The bacterium *Caulobacter crescentus* better survives exposure to high salinity if it is first exposed to a lower concentration of salt (Mathis and Ackermann, 2016). In this sense, probably any organism with second order metabolic processes can be shown to possess a simple faculty of memory.

In more complex organisms, specifically in animals with complex nervous systems, memory faculties are sufficiently developed that they must be differentiated into sub-categories. In humans, these are described as sensory memory, short-term memory, and long-term memory. Sensory memory refers to the ability to retain impressions of information received by a sense organ after the initial stimulus has ceased. Each sense modality can independently store traces of stimuli as sensory memories. Usually operative for periods of one second or less, it is the form of memory associated with perception, as stimuli which enter this memory buffer are retained long enough to be integrated into the perceptual schema (Pastukhov, 2016).

Short-term memory, also called working memory, is a memory buffer which enables the cognitive processing of sensory information. With a temporal range of less than one minute, and a storage capacity of less than ten items of information, considered across sensory modalities, it is deeply integrated with the initial cognitive processing of

the external environment (Jonides et al., 2008).

Long-term memory refers to the storage of relevant information over extended periods of time. This information can be stored for an indefinite period, and long-term memory systems are characterized by their large capacity. This form of memory is closely connected with the ability to recall information using associative networks (Blumenfeld and Ranganath, 2007).

Long-term memory, at least in humans, can be characterized as implicit or explicit. Implicit memory, also termed procedural memory, is the form of long-term memory responsible for the ability to reproduce complex tasks, and to acquire skills through repetition. Explicit memory is the storage of specific pieces of information which can be recalled when necessary. In humans, this form of memory is closely associated with phenomenal experience, and the entry of the recalled information into the field of consciousness (Rugg et al., 1998).

Explicit memory can be divided into semantic and episodic memory. Semantic memory enables the retention of items of knowledge about the world, such as concepts and facts. Episodic memory, the subject of this paper, is the faculty enabling an organism to retain information about specific previously encountered events, their relation to one-another, and their spatial and temporal contexts. The episodic memory trace for any given event must therefore contain at minimum three pieces of information: what the event was, when it happened, and where it happened. This can be abbreviated to “what, when, where” (WWW) information (Tulving, 1985).

## 1.2 The Structure of Episodic Memory

In humans, episodic memory is often called autobiographical memory, and is therefore closely associated with a concept of a *self* that ranges across time and space. The explicit conscious recall of temporal sequences of experienced events also closely ties episodic memory to the development of a concept of causation, in addition to reflection on past events, and projection of future possibilities, which necessitates a developed temporal representation divided into *past*, *present*, and *future*. There is however significant evidence that forms of episodic memory can exist in other species without these requirements.

Until comparatively recently, episodic memory was thought to be a distinctively human faculty, not dissociable from introspection, causal reasoning, the self-concept, and complex and explicit temporal representations allowing for mental time travel (Tulving, 1985). Observations suggesting the presence of something resembling episodic memory in other animals were dismissed as complex forms of conditioning and habituation, interacting with species-specific and hard-wired neurological architecture (Petri and Mishkin, 1994).

It is important to make a careful distinction between episodic memory and conditioning, especially when discussing the presence of behaviors seemingly requiring a system similar to episodic memory in non-human animals. Conditioning involves the acquisition of implicit memory. In some circumstances, conditioned responses may give the illusion of the involvement of episodic memory (Petri and Mishkin, 1994).

In both classical and operant conditioning, implicit memory is necessary to recall

the association between a stimulus and a response, and the temporal persistence of this memory can be impressively extended, but in neither case is any further faculty required. There is no need to code each stimulus response pair as an event occurring at a time and place in order for conditioning to be effective. The animal only needs to be able to give the learned response to the stimulus, without regard to the spatiotemporal context of the initial event.

Because of the ability of many animals to retain an impressive number of long chains of stimulus-response pairs, complex examples of conditioning can easily mimic forms of episodic memory. For example, a dog does not need to remember when a person left in order to anticipate their time of return. The ability to pair a specific sound, for example that of a train passing at a set time, with the subsequent arrival of the person, is sufficient.

While this is an important distinction to keep in mind, many critics of the existence of episodic memory in non-human animals have gone to great lengths to explain behaviors in ways that enable avoidance of attribution of episodic memory, which was historically seen to require faculties beyond those that could be reasonably expected of these animals (Tulving, 1985). The necessity of the existence of faculties resembling episodic memory to explain the cognitive abilities of many different species is now generally accepted. This form of memory has been termed “episodic-like memory,” in order to emphasize that it does not necessarily entail the same phenomenological correlates and conceptual apparatus as human episodic memory (Clayton and Dickinson, 2007).

### 1.3 The Minimal Requirements for Episodic-like Memory in Animals

Given that episodic-like memory in animals does not necessarily have the same structural, conceptual, or phenomenological instantiation as human episodic memory, although certainly allowing that it could, the necessary conditions for any form of memory able to encode WWW information must be discussed. This will be followed by a description of examples of episodic-like memory in various species, which will show that three different forms of this faculty can be distinguished. These forms enable increasing degrees of behavioral complexity and flexibility, and while the transition between the first two can be explained in continuous terms, there is a deep discontinuity between these two and the third form.

Keeping to the conceptual level of understanding, the neural substrates of episodic-like memory must be able to record specific distinguishing information about the event itself, *what*, information placing the event in a temporal context relative to other events stored, and a record of the duration of the event, *when*, and information on its physical location or cognitive context in relation to other events, *where*. As specified by Montemayor, this requires that the organism have the ability to measure time using both cyclical and durational clocks. The cyclical clock can be entrained to the circadian cycle, or to another reliably repeating physical process, such as the solar year, and the durational clock can take the form of a “stopwatch” that can be triggered and stopped at the start and end points of the event (Montemayor, 2013). This system of clocks and event tags yields a matrix of events arranged temporally and qualitatively in relation to one another.

It may be questioned why a perpetually updating record of the total time elapsed

from the occurrence of the event until the present, as might seem necessary to yield an episodic memory matrix according with the phenomenological experience of humans, has not been included. This omission is deliberate. Concepts of *past* and *present* are not necessary for the construction of an adequate system that preserves the empirically observable features of episodic memory, once the requirements that such a system include conceptual apparatus to account for ideas of causality and the self, are relaxed (Montemayor, 2013). The externally observed behavioral capacities necessitating the presence of simpler forms of episodic memory can be sufficiently accounted for by the above treatment, and will now be described in greater detail. These forms of episodic memory will collectively be described as *retrospective*. As the term implies, they are exclusively concerned with the encoding of WWW information on past events, with the caution that the use of the term *past* here should not be taken to imply the necessary existence of a concept of *the past* in these animals.

While the terms *retrospective* and *prospective* occur in the literature in conjunction with discussions of episodic memory, there is no general consensus on their use. Some authors divide all explicit memory into retrospective and prospective memory, in an alternative taxonomy to the declarative and episodic distinction. Some use the term “prospective memory” to describe the specific ability to “remember to remember,” and others use the term “prospective episodic memory,” but with widely varying definitions of the relationship between this form of memory, planning, and visualization (Brandimonte, Einstein, and McDaniel, 1996; Graff and Utzl, 2001). This paper will give a specific definition of prospective episodic memory, and show how it requires

qualitatively distinct cognitive abilities from the retrospective form.

#### **1.4      Retrospective Episodic-like Memory in Animals**

As described above, the presence of episodic memory in non-human animals has traditionally been denied. Their abilities have been thought to be wholly explicable through a combination of conditioning and familiarity with objects and situations, leading to the description of animals as being “stuck in time,” or living in a “continuous present” (Roberts, 2002). In addition to the behaviorist conceptual schema of the mid-twentieth century, this denial can be accounted for by a rigid conception of episodic memory as necessarily explicit. This commitment to the explicit quality of episodic memory also led to the argument that its development in humans was dependent upon the acquisition of syntactic language, in order to express the required conceptual apparatus in propositional terms (Tulving, 2002).

Since the early 1990s, empirical evidence has accumulated suggesting the existence of forms of episodic-like memory in many different species with complex nervous systems. This evidence ranges over a wide variety of species in disparate clades, and suggests that episodic-like memory, being an extremely useful faculty, has evolved independently multiple times. For example, forms are found in bees, cephalopods, and vertebrates (Pahl et al., 2007; Jozet-Alves, 2013). Within vertebrates, it is found in both birds and mammals. This is likely a case of convergent evolution, given the distant common ancestor of these two groups. This repeated evolution shows that the systems underlying episodic-like memory are not so rigid as to develop only over long time periods under very specific selection pressures, instead being a flexible faculty

instantiated easily by a sufficiently complex nervous system.

The specific forms of episodic-like memory found differ in each case, being tailored to the individual requirements of a species, but when taken together can be arranged along a continuum from extremely *domain-specific*, to *domain-flexible*. Domain-flexible episodic-like memory can develop smoothly from more domain-specific forms, as the background requirements for WWW information storage are the same in both cases.

For example, bees have a sophisticated ability to encode time, place, and object information regarding the location and nectar quality of flowers, and their connections to the circadian cycle, but cannot apply this ability flexibly to other domains, such as those involving objects other than flowers. Their form of episodic-like memory is tightly bound to its specific evolved application, and they are unable to generalize.

Rats, cuttlefish, pigeons, and pigs have shown a retrospective episodic-like memory with respect to aversive experiences, foraging behavior, and food rewards in specific situations, also distinguishing between preferred and less sought after food items (Kart-Teke, Silva, Huston, and Dere, 2006; Jozet-Alves, 2013; Zentall, Singer, and Stagnar, 2008; Kouwenberg, Walsh, Morgan, and Martin, 2009). This is different from simple conditioning or associative learning, as they are able to make active choices of where to go and what to do based on their past experiences. Their form of episodic-like memory is not tightly bound to a circadian cycle, as they display it in the absence of circadian cues (Jozet-Alvez, 2013), show long term information retention, and have some ability to generalize across domains, with respect to past events. They can also display

the faculty in experimental environments unlike those they would encounter in the wild, and in novel situations (Kart-Teke et al., 2006).

All of these examples of episodic-like memory are retrospective. They enable the animal to modify behavior based on prior events in ways that necessitate a system encoding WWW information regarding these events. As described above, the outputs of durational and interval clocks, coupled with a system of event tags and a representational matrix specific to the requirements of the species in question, are sufficient to account for their expression (Montemayor, 2013). Some species exhibit more domain and context specific forms, and others show a more flexible and generalizable faculty, but the cognitive demands between these two differ quantitatively, and in terms of the degree of integration between cognitive domains devoted to processing different aspects of the animal's environment. Flexibility can develop smoothly from specificity, and causal reasoning, a self-concept, and explicit temporal representations are not required.

### 1.5 Prospective Episodic-like Memory in Animals

There is another form of episodic-like memory, shown by a small group of species, that is not explicable without the introduction of additional conceptual requirements beyond those necessary for the above examples. This involves the projection of the WWW matrix into a conceptual space where it can be populated by hypothetical events, enabling the animal to leverage the abilities granted by episodic-like memory to develop the capacity to plan future actions based on hypothetical scenarios. This can be described as a *prospective episodic memory* (PEM), a seemingly paradoxical use of memory to simulate future events. Humans display this capacity constantly, and

our dependence on future planning is responsible for the assumptions we often make regarding the conceptual requirements for episodic memory in general. Other animals showing PEM include great apes and some corvids. In addition to predicting future events, they can decide which actions to take to increase the likelihood of desired future outcomes.

Scrub jays depend on caching foods in order to survive periods of scarcity. Squirrels also show this behavior, but in contrast to squirrels, which have limited ability to remember the specific locations of the nuts they cache, relying more on a combination of memory of the general area coupled with a keen sense of smell to relocate their stored items, scrub jays can remember the specific locations of many individual food items, the elapsed time since they cached them, and the respective decomposition rates of multiple food types (Clayton and Dickinson, 1998). The birds use this information to determine the best order to retrieve their foods. For example, a mealworm decomposes faster than a peanut, so a jay will make sure to return to the mealworm first, if they cached both items on the same day. These observations show that scrub jays possess an extremely well developed retrospective episodic-like memory.

However, in addition to these abilities, scrub jays will cache foods in areas where they know there is generally little to eat, but where they must later go for some other purpose. They will also cache specific preferred food items, such as mealworms, in areas where these particular foods will not later be available. Similar examples of PEM are exhibited by other corvids (Salwiczek, Watanabe, and Clayton, 2010).

Great apes, and possibly other primates, also possess PEM, although possibly not

to quite the same degree of sophistication as corvids (Martin-Ordas, Huan, Colmenares, and Call, 2010). Bonobos and chimpanzees will preferentially seek out frozen juice over grapes, when both foods are hidden in specific areas, if they have waited for five minutes before being given access to both areas. If they have to wait an hour, they do not bother to search for the frozen juice, knowing that it has melted and become unobtainable. Instead they immediately search for the grapes (Martin-Ordas et al., 2010). PEM may also be hypothesized to exist in other intelligent and highly social animals, such as elephants, parrots, and cetaceans.

### **1.6 Prospective Episodic Memory, Time, and The Agent Representation**

The development of PEM is closely connected to the evolution of temporal representation, and cannot develop smoothly from REM for several reasons. Foremost among these are the necessity of a second-order cognitive simulation of an episodic memory matrix, and the associated requirement for some minimal form of hypothetical reasoning enabling different scenarios to be run on this simulation. In order to differentiate the first-order and simulated matrices, a form of temporal representation that distinguishes *past* from *future* is necessary, noting that this need not imply the existence of these as explicit concepts. The hypothetical scenarios explored must also be grounded in reference to a minimal stable representation of an agent, which is not necessarily equivalent to a self-concept. Additionally, the simulated *future* scenarios must be continuously updated and modified with the accumulation of events in the *past* matrix. A system of event tags and clock outputs is inadequate to account for these features, the origin of which must instead be sought by examining other cognitive capacities and

behavioral systems shared by the group of species that are capable of PEM.

What all species thus far shown to possess PEM do share is a high degree of social integration. Corvids, and humans and other apes, all live in immersive social worlds and are tightly associated with specific other individuals in cohesive groups. Their capacity for episodic memory is necessary for them to keep track of the status of their social relationships, by remembering specific events involving interactions with other members of their group. Additionally, they all show behaviors that necessitate some degree of prediction of the belief states of other individuals based on their observations of events experienced by these individuals, and in some cases use this predictive ability to manipulate these belief states to achieve desired outcomes.

For example, scrub jays will re-cache food in a different location if observed by other individuals while caching, and differ in their reactions according to which other bird is observing (Dally, Emery, and Clayton, 2010). Jays are more likely to re-cache food if observed by a socially dominant bird with whom they do not have a close relationship, than they are if viewed by subordinates or partners (Dally et al., 2010). Crucially, they do not show this behavior until they have themselves been stolen from; it does not appear to be completely innate in its full expression. Their reactions to being observed are also matched to specific individuals in their social group, and remain stable over time (Dally et al., 2010).

Non-human apes are well known to engage in a variety of behaviors suggesting an ability to attribute mental states to other individuals, including elaborate and long term alliances, deceptions, and “political” behavior (Whiten, 2001). Humans take these same

characteristics to an extreme level, building elaborate networks of empathy, and webs of deception, in order to maintain their intricate social networks and complex societies. Taken together, these observations imply that all these species share certain components of a theory of mind.

### **Empathy**

When I confront a human being as my Thou and speak the basic word I-Thou to him, then he is no thing among things nor does he consist of things. He is no longer He or She, a dot in the world grid of space and time, nor a condition to be experienced and described, a loose bundle of named qualities. Neighborless and seamless, he is Thou and fills the firmament. Not as if there were nothing but he; but everything else lives in his light.

— Martin Buber, *I and Thou*

## **2.1 Prospective Episodic Memory and Theory of Mind**

“Theory of mind” is a term generally used in the fields of developmental and animal psychology (Premack and Woodruff, 1978). In humans, it refers to the ability of a person to both gain a second-order awareness of their own mental states, which can be termed metacognition, and to theorize both that other individuals have mental states, and that the structure of their belief states and emotional systems is constructed in a similar way to their own. This enables the prediction of the behaviors of others based on a model of the operations of their minds, and the successful imputation of their cognitive and affective states, a faculty which is termed mind-reading (Premack and Woodruff, 1978).

In general terms, metacognition is associated with, and likely necessary for, introspection, while mind-reading shares a similar relation to empathy. The full development of these abilities, as is seen in most adult humans, can be referred to as a

“complete theory of mind,” although the degree to which an individual can engage in metacognition, and correctly infer the mental states of others, are both extremely variable (Mitchell, 2011). The extent to which the first is expressed varies according to the psychological resources applied to second-order cognition, while the maximal expression of the second depends on skilled observation and background knowledge of contextual information relevant to particular individuals (Mitchell, 2011). While this paper will not be concerned with a detailed analysis of the variation in expression of the components of theory of mind in adult humans, specific cases affecting their development will be briefly described, as relevant to the central arguments of the paper.

Similarly to episodic memory, the development of theory of mind has traditionally been thought to require the presence of a self-concept, based on the observed relation between the two in humans (McHugh and Stewart, 2012). As in the case of episodic memory, there is reason to argue that the components of theory of mind can be dissociated, but that this dissociation can only occur in one direction. The dissociable components of a complete theory of mind will now be explored individually in greater detail.

## **2.2 The Components of Theory of Mind**

As described above, mind-reading can be minimally described as the ability of an organism to successfully model the affective and/or cognitive state of another individual. These two categories of mind-reading can be described as forms of empathy. Affective empathy is the ability of an organism to correctly identify, and often mirror, the affective states of another individual (Perry and Shamay-Tsoory, 2013). This faculty requires the

least cognitive complexity of the different components of mind-reading, and as will be described below, there is much empirical evidence that the evolution of affective empathy long precedes that of cognitive empathy.

Cognitive empathy is the ability to impute the cognitive states of others based on knowledge of their prior and current external environmental context (Perry and Shamay-Tsoory, 2013). Both of these faculties can be accounted for without the need to invoke the presence of a self-concept. They do require a cognitive model of an *agent*, constructed according to the psychology of the species in question, and an ability to generate a simulation of the sequence of events leading to the current context of this agent, but there is no need for the mind generating this agent model to be able to reflect upon its own mental states in order to account for the ability.

Metacognition refers to the ability of an organism to engage in second-order analysis of its own first-order mental states, to “think about thinking” (Terrace, 2006). This faculty is intimately related to the self-concept, as, in this case, the agent modeled by the organism must be self-reflexive. The cognitive substrate being analyzed must necessarily be the first-order cognition of the organism itself. Metacognition can thus not be dissociated from a self-concept of some form.

As complete theory of mind requires both mind-reading and metacognition, an account of the precise relationship between these two will now be given. This will also address the question of whether the full development of both can be wholly explained by a quantitative increase in the complexity and flexibility of the faculty of theory of mind taken in itself, or whether it requires the importation of abilities deriving from quite

different cognitive domains.

### 2.3 The Empathy-first Model

The traditional position on the development of theory of mind is that mind-reading ability is a consequence of the projection of metacognitive abilities onto others, that we learn to mind read by realizing that others have minds like our own, and therefore that metacognition comes first. This is known as the “simulation theory” (Gallese and Goldman, 1998). The alternative, and more recently developed, “mind-reading first” theory of the origin of theory of mind argues that mind-reading developed before metacognition (Carruthers, 2009). This makes sense from an evolutionary perspective, as mind-reading has much more obvious immediate utility than a faculty of introspection. The mind-reading first theory, which I will here term *empathy-first* argues that our perceived knowledge of the contents of our own mind, our sense of privileged access to our own mental content, results from us turning our mind-reading capacity inward and applying our other-directed theory of mind to ourselves. This theory thus denies the existence of introspection unmediated by a prior act of swift, unconscious self-interpretation (Carruthers, 2009).

As will be shown below, there is convincing experimental and observational evidence that affective empathy appears extremely early in evolutionary terms, and several species are capable of at least some degree of cognitive empathy. Given that neither faculty requires the presence of a self-concept, this data suggests that self-concept, and thus metacognition, is not a necessary condition for mind-reading. This is supported by further observations that very few species have given evidence of

metacognition, with the concession that finding and testing for empirical correlates of this faculty is extremely difficult.

#### **2.4 Empathy across Phylogeny**

One of the most basic forms of affective empathy, emotional contagion, has been observed in a great number of unrelated species, including such seemingly unlikely examples as crustaceans. Woodlice have been shown to exert calming effects on their excited neighbors, enabling them to remain in cohesive groups (Broly and Deneubourg, 2015). Similar examples have been seen in other arthropods, and emotional contagion is widely distributed among vertebrates.

Such a simple ability is certainly not comparable to a developed form of mind-reading, but there is evidence that more complex forms of affective empathy can develop from it in a smooth fashion. For example, an animal can, instead of simply mirroring the emotional state of another, modify its behavior in order to reinforce or change that state, based on its own positive or negative emotional responses. This behavior has been well documented in many vertebrates, including iguanas and chickens, and in a more complex form in rats. Juvenile iguanas will engage in soothing activities such as head rubbing and body contact with agitated group members, in order to maximize group cohesion before setting off on journeys, and hens respond with increased heart rate and stress behaviors when they observe puffs of air blown at chicks, followed by extended brooding behavior (Alberts, Carter, Hayes, and Martins, 2004; Edgar, Lowe, Paul, and Nicol, 2011).

Rodents display a more developed form of empathy. Multiple studies have shown that a rat will act to assist a distressed companion, even when given the choice between

consuming an entire piece of chocolate themselves, or sharing it with the other rat. In one example, a rat was given the opportunity to choose between three possible courses of action after encountering a pile of chocolate adjacent to another rat in a small cage. They could release the other rat from the cage before sharing the chocolate, eat the chocolate first, and then release the second rat, or simply eat all the chocolate and leave, leaving the other rat caged. A large majority of rats released their companion and shared the chocolate, even if they were prevented from engaging in social behavior afterwards (Bartall, Decetyl, and Mason, 2011). The second study was similarly constructed, involving one rat being given the opportunity to extricate another from a container where it had to stand in shallow water. The majority of rats aided their companion, and the percentage was increased when they themselves had been in the water in a previous trial, showing that affective empathy was increased by their individual experience of a similar predicament (Sato et al., 2015). Similar behavior has been seen in several other bird and mammal species.

Cognitive empathy is less widely distributed, but, as described in the prior discussion of episodic memory, has been shown to exist in corvids and some primates (Dally et al., 2010; Whiten, 2001), and may be hypothesized to exist in certain other groups showing complex social behavior, such as elephants, parrots, and cetaceans. An important point to note here is that retrospective episodic-like memory is widely distributed across species that display affective empathy only, but all species so far shown to possess prospective episodic-like memory also exhibit cognitive empathy.

## **2.5 The Difficulty of Testing for Metacognition**

Tests for metacognition in non-human animals are difficult to perform, and hard to distinguish from more generalized tests for components of self-concept, which themselves often give confounding results. The classic example is the mirror test, where a paint spot is applied to an animal without their knowledge, and in an area that can only be seen by examining their reflection in a mirror (Gallup, 1970). Animals that try to remove or otherwise orient toward the spot using information gained from their reflected image are said to pass the test, which is used as a proxy for self-awareness.

The results from this test are extremely uneven. Some results are in line with general expectations, for example, elephants, apes, dolphins, and magpies sometimes pass (Plotnik, de Waal, and Reiss, 2006; Westergaard and Hyatt, 1994; Reiss and Marino, 2001; Prior, Shwarz, and Güntürkün, 2008). However, so have pigeons, manta rays, and ants (Uchino and Watanabe, 2014; Ari and D'Agostino, 2016; Cammaerts and Cammaerts, 2015). Also, parrots, and corvids other than magpies, do not pass the test, and among the species that do, there is great individual variation. Some animals only pass after various forms of training, such as having the spot applied to a visible part of their body the first time, or with the use of food rewards. While the mirror test is certainly testing for something, it is mistaken to think of it as a clean test of the presence or absence of a self-concept.

Other tests specifically for metacognition have also given interesting but inconclusive results. For example, in an ingenious study, monkeys were trained to complete tasks requiring that they remember visual images, including the opportunity to indicate that they were uncertain of whether they had seen the image before. When the

monkeys had their memories of certain images degraded by transcranial magnetic simulation, they were more likely to respond that they were uncertain, giving some evidence of metacognitive abilities (Washburn, Gullede, Beran, and Smith, 2010).

As the above examples show, attempts to demonstrate metacognition in animals have given some suggestive yet generally unconvincing evidence. In contrast, and as described earlier in this paper, there are many experimental and observational studies quite conclusively demonstrating the presence of mind-reading abilities in several species. A concession that must be given here is that although both apes and some birds, particularly corvids, do show conclusive evidence of mind-reading in a social context, they do not pass the gold-standard test for mind-reading, the false-belief test, in which an observer must realize that another individual has a belief about the location of an object that differs from their own, after they observe the object being moved without knowledge of the other individual (Wimmer and Perner, 1983).

Human children do not perform well on this test until the age of four, and apes in general have fared poorly (Krachun, Carpenter, Call, and Tomasello. 2009). Scrub-jays will conceal inedible items while being observed in order to deceive possible thieves (Dally and Clayton 2010), but this strategy does not require a comparable level of insight. It is probable that reliably passing the false-belief test requires the ability to generate specific propositional statements, and is therefore dependent on linguistic abilities unique to humans, and that failure to pass this test has surprisingly little impact on general mind-reading skills. This is supported by the observation that, in certain unusual cases of human language development, mind-reading as a social ability is dissociable from the

ability to pass the false-belief test. One of the best examples of this is the case of Nicaraguan Sign Language.

In the late 1970s, a group of fifty Nicaraguan deaf children were placed in a special school in San Judas; due to the dislocations of the Sandinista revolution they were not effectively educated in signing, and remained linguistically isolated from their teachers. By the early 1980s there were 400 children at the school, and they progressively developed their own sign language, now known as Nicaraguan Sign Language, or NSL. There are now approximately three thousand speakers of NSL, some of them third generation. The first generation form of NSL did not possess words to clearly demarcate the concept of false beliefs, and first generation speakers do poorly on false belief tests (Morgan and Kegl, 2006). Although the next two generations of speakers have developed a more diverse vocabulary, which does express this concept efficiently, and have no trouble, the first generation speakers never perform quite as well, even when provided with this improved vocabulary (Morgan and Kegl, 2006). Although they perform poorly on the test, and in certain specific domains of mind-reading, they clearly have little difficulty navigating their daily lives in human society, which requires mind-reading on a frequent basis.

Taken together, the above evidence gives support to the argument that mind-reading can and does exist independently of metacognition, and that there is a continuum of degrees of the ability to mind-read that a species or individual can possess, conditioned by the degree to which both affective and cognitive empathy are developed, in addition to the presence or absence of different degrees of linguistic capability and thus the ability to

express propositional statements.

## **2.6 Mind Reading and Prospective Episodic Memory**

A strong correlation between the possession of a well developed faculties of mind-reading and of prospective episodic memory has been demonstrated, but retrospective episodic memory is widely distributed among species which are poor mind-readers. This shows that retrospective episodic memory can exist prior to, and is entirely dissociable from, mind-reading taken in the full sense as requiring both affective and cognitive empathy. The directionality of the relationship between mind-reading and prospective episodic memory must now be examined.

This paper proposes that PEM cannot develop smoothly from REM without the prior development of a basic form of the mind-reading component of theory of mind, here referred to as *minimal theory of mind*, and that the development of metacognition requires the prior presence of all three of these faculties. This is supported by evidence from animal behavior, neuroscience, and developmental psychology.

Several species show evidence of mind-reading without significant evidence of PEM. These include dogs, pigs, and goats. Dogs show mind-reading capabilities, at least with respect to humans, surpassing those of non-human primates, which can be partially explained by the intense selection for this ability due to domestication. They are more likely to attempt to steal food when a human is facing away from them, or when they can see that the human's eyes are averted (Maginnity and Grace, 2014). Dogs also use attention getting behaviors to induce other dogs to make eye contact before initiating play (Horowitz, 2008). Pigs can learn how to follow mazes to receive a food reward by

observing the path taken by another pig (Held, Mendl, Devereux, and Byrne, 2001), and socially subordinate goats preferentially seek out food which dominant animals cannot see (Kaminski and Tomasello, 2006).

There is evidence from the neuroscience and developmental psychology literature that, at least in humans, while mind-reading develops independently of episodic memory, and can exist without it, the reverse is not possible. Multiple studies have shown that brain lesions impairing episodic memory and future planning abilities do not affect mind-reading. For example, a woman with severely impaired episodic memory development due to a hippocampal lesion sustained in early childhood, performed normally on all tests of mind-reading, and was indistinguishable from controls (Rabin, Braverman, Gilboa, Stuss, and Rosenbaum, 2012). Studies of child development give similar results. A series of experiments with children between the ages of three and six demonstrated that mind-reading developed both prior to and independently of the development of episodic memory.

There is as yet no evidence of dissociation in the opposite direction, that is, of well developed prospective episodic memory without significant mind-reading ability, in either humans or other animals. Several workers have assumed that such cases must exist (Rabin et al., 2012), but as of yet have not found actual examples. It is the contention of this paper that dissociation in this direction is not possible; that is, that the mind-reading component of theory of mind is necessary for the full development of episodic memory in humans. Additionally, it is argued that this component of theory of mind must come prior to PEM evolutionarily, and that animals that show PEM acquired the ability due to an

evolutionary pathway that was opened by the previous development of mind-reading.

## **2.7 Prospective Episodic Memory and Metacognition**

Given the convincing evidence that mind-reading is indeed necessary for the development of PEM, and thus that the relation REM → MR → PEM has been established, the next and final assertion of this paper, that PEM is necessary for metacognition, must be supported. As described above, the experimental evidence for metacognition in animals is inconclusive, but many of the animals that have shown some suggestions of metacognition, such as corvids and primates, do possess PEM.

There are also many animals that exhibit PEM without any evidence of metacognition, such as corvids other than magpies, and certain monkeys, and within species that do show some suggestions of the latter faculty, the ability, if it does exist, is very unevenly distributed among individuals. For example, even in chimps, which pass the mirror test more reliably than any other species, only 75% of young adults succeed, and the proportion is far less in both younger and older individuals (Povinelli, Rulf, Landau, and Bierschwale, 1993), even though evidence of PEM is much more reliably found.

In the few species that occasionally pass the mirror test without obvious evidence of PEM, there is either no replication of the experiments, as in ants and manta rays, or the ability is only shown after extensive training and shows some resemblance to conditioning, as in pigeons (Uchino and Watanabe, 2014). Humans are the only animals that are definitively known to have a self-concept, and to engage in metacognition. They also show extremely well developed PEM. The above observations give evidence that

PEM is required for the development of metacognition.

## **2.8 Prospective Episodic Memory and Metacognition in Human Evolution**

Further evidence that PEM is necessary for metacognition can be seen in the sequence of appearance of the correlates of modern human mental characteristics in human evolution, specifically the timeline and distribution of the production of art among hominins. This evidence also supports the well known argument that language ability is a necessary condition for fully developed metacognition. It can be established, using phylogenetic bracketing, that hominins subsequent to the most recent common ancestor of chimps and humans possessed at least a basic form of PEM. Examination of whether some of them also possessed metacognitive abilities requires a short excursion into paleoanthropology.

The extensive production of artworks is only seen among anatomically, and presumably neurologically, modern humans, with some suggestions among neanderthals, and the recently discovered denisovans, such as a collection of inconclusive evidence of the use of feather accessories and cosmetics among neanderthals (Finlayson et al., 2012), and a finely worked green chlorite bracelet possibly crafted by denisovans (Gibbons, 2011). As is well known, neanderthals also sometimes buried their dead (Rendu et al., 2014), and use of red ochre, possibly as body paint, was ubiquitous among archaic hominins, including neanderthals (Roebroeksa et al., 2012).

Importantly, the artistic activities of these archaic hominins are self-reflexive in a narrow sense, involving the possible production of jewelry, cosmetics, body paint, and body adornments such as feather headdresses, along with occasional and disputed

evidence of grave goods including animal bones, horns, and stone tools. This suggests that they did possess a nascent self-concept with reference to their bodies in their social context, but that they did not necessarily possess the same form of metacognition found in modern humans, who produce massive quantities of symbolic and figurative art, enabled by extensive use of metaphor and abstraction. Given that they also lived in small, isolated groups (Conrad and Richter, 2011), and did not generally trade in symbolic artifacts, as evidenced by their tools being generally practical, and produced mainly from materials found locally (Conrad and Richter, 2011), their awareness of self was also possibly less integrated to their social environment than it is among modern humans, possibly implying that their linguistic ability was less flexible than that of modern humans when applied to description of social situations.

The above example provides evidence that a form of self-concept in a species known to exhibit PEM can originate prior to the development of full metacognitive ability. This paper proposes that full metacognition requires that this nascent self-concept be integrated with the mind-reading component of theory of mind, leading to a state where mind-reading abilities are hooked up to the concept of self, and a full theory of mind, entirely available to conscious attention, is present. As will be described below, this full metacognitive ability, coupled with a flexible and tensed language, then enables the origin of metaphor and symbolism, and artwork separated from immediate reference to its individual creator. The appearance of this behavioral modernity in humans occurred both suddenly and seemingly fully formed, and long after the evolution of physical modernity (Zilhão 2007), which suggests that large scale modifications of neural

architecture are not required, and would support the hypothesis that the integration of the early self-concept with mind-reading was a comparatively sudden event, which could perhaps have occurred in individual groups.

If true, this would have been extremely disruptive to the social architecture of these human groups, and may be correlated with the appearance and sudden dispersal of behaviorally modern humans from Africa and the Levant roughly 70 000 years ago. Groups of modern humans reached Australia by 60 000 years ago, and, by 40 000 years ago, all of Eurasia had been populated (Beyin, 2011), with the extinction of neanderthals in Europe occurring by 30 000 years ago (Higham et al., 2014).

## **2.9 Metacognition and Behavioral Modernity in Humans**

Finely made figurative artwork, in the form of cave paintings and portable sculptures, appears fully developed in the archaeological record, without significant evidence of a slow incremental development (Sadiera et al., 2012). The extensive production of this artwork can be argued to be a proxy for metacognitive ability, as figurative artwork by necessity involves symbolism and metaphor, and requires the viewer to be able to place themselves into a cognitive state where the image evokes its referent, requiring second-order cognition. Paleolithic cave paintings are also saturated with narrative. For example, in Chauvet Cave in the Ardeche region of France, groups of painted animals are positioned around gurgling water sources to simulate herds at water holes, and outlines are superposed to generate movement effects when illuminated by torchlight (Azéma and Rivière, 2012). This requires the creator or viewer to be able to bind their second-order cognitive states into a continuous flow, in which they place

themselves in this simulation, again requiring well developed metacognitive ability connected to skilled visualization of the referents of the images, and the ability to identify the self with the visualized observer.

Modern humans are also the only hominin to show extensive trade networks and the production of tools that require incremental planning in their manufacturing process. The differences in the methods of tool production and forms of technology between neanderthals and modern humans are intriguing. Both species produced sophisticated stone tools, these being flakes in the case of neanderthals and blades in that of early behaviorally modern humans. Neanderthals, and some archaic anatomically modern humans, produced their flakes by a method known as the Levallois technique (Wynn and Coolidge 2004). This entailed the skilled preparation of a prismatic core from flint, obsidian, or other suitable fine-grained stone, using a hard stone hammer. For the production of large rounded flakes for knives and ax-heads, a “tortoise core” was sometimes crafted. Once this core, the shape of which was tailored specifically to the type of tool being produced, had been created, a striking platform was made through the careful removal of small flakes using pressure from a sharp object (Wynn and Coolidge 2004). The shape of the core and platform was determined by knowledge of the fracturing characteristics of the type of stone being used, and the type of tool required. At this point, the tool was struck from the platform by a single blow with a bone or antler hammer, giving a single flake with a uniform sharp edge.

This method was used to manufacture spear points, scrapers, boring tools, ax-heads, and knives, among others, and had the advantage that multiple tools could be

struck from one prismatic core, which could be carried so that a fresh sharp tool could be struck from the core on demand (Wynn and Coolidge 2004). Neanderthals also developed a simple form of furnace where birch bark was burned in the absence of oxygen to generate pitch, which was used as a glue to bind ax-heads to wooden handles, and stone points to wooden spears (Kozowyk, Soressi, Pomstra, and Langejans, 2017) among other similarly skilled technologies, including possible use of manganese dioxide as a fire starter (Heyes et al., 2016).

Early modern humans, in contrast, created a greater diversity of tools, using a method similar to that employed when carving a statue. Small pieces of stone would be progressively removed from the core, slowly changing its shape into the form of the finished tool, the quality of which would vary according to the skill of the maker, and which could be repeatedly retouched (Ambrose 2010). This method enabled the production of tools specifically suited to many different intended uses, and had the advantage of not requiring the degree of procedural memory and hand-eye coordination needed by the neanderthals, who practiced and refined the Levallois method from early childhood (Stapert, 2007), but had the disadvantages of being more wasteful of stone, not enabling the use of a prepared core to create multiple tools on demand, and not reliably producing such uniform and sharp edges. Even now, humans have trouble acquiring skill equivalent to neanderthals in the Levallois technique (Wynn and Coolidge, 2004). Early modern human technology also changed quickly over time, whereas neanderthal methods and products were extremely stable over tens of millennia (Wynn and Coolidge, 2004).

A hypothesis can be proposed that neanderthal technology was dependent on the

acquisition of sufficient experience in a given domain to be able to enter a form of flow state, characterized by excellent procedural memory and consummate skill, but unconnected to the same form of inner dialogue occurring in modern humans, and was taught to children by accomplished masters. This is intriguingly similar to states of mindfulness induced through meditative practices. In the Levallois method as practiced by neanderthals, the core produced bore little resemblance to the finished flake, and the quality of the tool was dependent upon a single, perfectly calibrated strike. If this failed, the core and striking platform had to be reworked from the start.

For modern human blade technology, the process of creation was accompanied throughout by visualization of the finished product, which was progressively sculpted from the stone. This process demanded that the creator be able to hold the image of the finished tool in mind as they worked. This method is more tolerant of error, but requires metacognition in order for the creator to maintain second-order cognition of the finished tool in addition to the first-order cognition required to expertly flake the stone.

## **2.10 From Retrospective Episodic Memory to Metacognition**

It is now necessary to provide a causal hypothesis explaining the empirically suggested evolutionary and developmental sequences described above. To reiterate, the discussion of metacognition gives evidence that the fully developed faculty is dependent on an integration between a nascent self-concept and the mind-reading component of theory of mind, and its full expression in modern humans is recorded by the origination of artwork and technology requiring an understanding of metaphor and symbolism, and the ability to sustain an internal dialogue. This hypothesis must now be integrated with

the analysis of PEM to show that the only way this early self-concept can develop is through the leveraging of a cognitive tool enabling PEM, which originated with mind-reading.

As established at the start of the paper, PEM, while not requiring a self-concept, does require a minimal concept of an agent, in order to provide a placeholder upon which to run simulations of the projected episodic memory matrix. REM does not require this, as it is not concerned with simulation, and therefore PEM cannot develop smoothly from REM without the importation or direct origination of this minimal agent concept.

The central hypothesis of this paper is that this agent substrate on which simulations necessary for PEM are run is directly imported from the agent model required by the full development of the faculty of mind-reading as expressed in cognitive empathy, and that mind-reading itself cannot develop without a preexisting faculty of REM, enabling the mind-reader to remember the sequence of events experienced by the individual they are modeling.

To put this argument in as simple terms as possible, retrospective episodic memory enables an organism to record prior sequences of events, using a WWW (what, when, where) matrix. The faculty of affective empathy, in its simplest form as emotional contagion, develops independently.

Cognitive empathy, and therefore the full expression of the mind-reading component of theory of mind, is enabled by the observation of other individuals, and their placement in the WWW matrix of REM with development of the ability to form minimal agent representations.

The WWW matrix is now duplicated, and this duplicated matrix is used to run simulations of future events, using the minimal agent representation originating from the observation of others, and previously placed in the WWW matrix, as a placeholder. This agent representation can now be used to predict the future behavior of other individuals, an extremely useful skill.

When cognized separately from the context of cognitive empathy, the future matrix with agent representation allows the individual to make future plans and modify their behavior based on the results of different simulations. This is now a fully developed faculty of PEM. With the development of increasing social complexity, and coupled with increasingly subtle forms of temporal representation, as will be discussed below, this minimal agent can develop into a nascent self-concept.

This early self-concept can now be directly hooked up to the faculty of mind-reading, correlating to the acquisition of syntactic and tensed language used to narrate social situations. The maximal integration of mind-reading and this basic self-concept enables the organism to engage in second-order cognition, yielding metacognition and internal dialogue. Metacognition then enables the understanding and production of symbolism and metaphor, and rapid and sustained technological change, as is observed in modern humans.

The advantage of this account is that it at no time requires the origination of a faculty that does not have immediate evolutionary benefit to the organism. For example, the development of mind-reading abilities is advantageous to any social animal, but a self-concept considered in itself seems of little immediate evolutionary value. Extensive

and empirically unsupported theoretical scaffolding would be required to explain the selective advantage of a self-concept as an isolated faculty. The theoretical structure proposed here is more parsimonious, and is capable of explaining the observed data in terms that allow the operation of well known and simple forms of selection pressure to generate massive complexity over geological time scales.

As in any evolutionary account of such systems, the aspects involving phenomenal experience cannot be directly treated here, although some light can be shed on the development of their conceptual and neurological correlates, which are fundamentally associated with the changes in temporal representation required to enable the organism to acquire each faculty.

### Time

So although we have no recollection of having existed before the body, we nevertheless sense that our mind, insofar as it involves the essence of the body under a form of eternity, is eternal, and that this aspect of its existence cannot be defined by time, that is, cannot be explicated through duration. Therefore, our mind can be said to endure, and its existence to be defined by a definite period of time, only to the extent that it involves the actual existence of the body, and it is only to that extent that it has the power to determine the existence of things by time and to conceive them from the point of view of duration.

– Baruch Spinoza, *Ethics*

### 3.1 Temporal Paradoxes, the Discrete, and the Continuous

As described above, each step of the progression REM → MR → PEM → MC requires elaboration of an organism's temporal representation system, and the integration of these changes into the systems of cognition and attention which are associated with phenomenal consciousness, at least in humans. The result of this process may be reflected

in the structure of the phenomenal experience of time. The occurrence of temporal paradoxes in human cognition may be a consequence of this progressive development, and show how the evolutionary process, as in many types of biological system, resulted in the formation of cognitive and representational structures that were satisfactory to enable functional cohesion and practical efficacy, but which were not constructed to optimize smoothness and consistency across all levels of analysis. An examination of these points leads to the conclusion that direct realism is false, at least with regard to the human experience of time, and that the way that we perceive time is fundamentally conditioned by faculties acquired for specific purposes throughout our evolutionary history.

The fundamental paradox of time perception, the first recorded examples of which are given in the conflict between Heraclitus and Parmenides, and in the paradoxes of Parmenides' student Zeno, is the irreconcilability of our perception of tensed time flowing from past, through present, to future, and our contradictory perception of the existence of an ordered sequence of events, existing in a static and tenseless time-block, which assume temporal and spatial coordinates relative to one another. This conflict has resulted in many examples of temporal paradoxes throughout history, some of which are deeply rooted in the structure of mathematics, physics, and other systems mapping onto our experience of the world.

Parmenides expresses the conflict in stark terms (Kirk, Raven, and Schofield, 1983). In order to accept temporal flow as real, we must claim that neither the past nor the future exist, and that we experience the existence of ourselves in the world only in the

present. The present is said to be the product of the past, which no longer exists, and the cause of the future, which does not yet exist. If time flows, then there is no stasis. If there is no stasis, then the present must be infinitely divisible. But if the present is infinitely divisible, then it is an infinitely small slice of existence positioned between non-existence on both sides. If the present is infinitely thin, then it is swallowed up by the past and the future, neither of which exist. So, in this system, there can be neither cause nor effect, and nothing can exist, given that something cannot come from nothing. But we experience change, so time flows, and we are back to the start.

In Zeno's arrow example (Kirk et al., 1983), the conflict is expressed in the irreconcilable yet extremely powerful perceptions of the arrow as moving continuously through space as time flows, and as occupying a single spatial position at any one instant. Attempts to defuse the paradox must contend that at least one of these two perceptions is false. Either time does not flow continuously from past to present to future, in which case the positions of the arrow at different time snapshots can be mapped as points in a static spatiotemporal coordinate system, or time does flow, the present is infinitely divisible, there was a past which no longer exists, there will be a future which does not yet exist, and our perception of the existence of a sequence of positions occupied by the arrow at different moments is false. Both of these positions can only resolve the paradox at the expense of invalidating a fundamental component of our perception of time.

The same paradox is expressed in multiple domains of mathematics. Calculus is often invoked as a solution to Zeno's paradoxes, but actually contains similar inconsistencies itself. As Bishop Berkeley famously pointed out, the concept of the

derivative as a ratio of infinitesimal quantities poses both logical and metaphysical problems. The addition of limits to the foundations of calculus was an attempt to ameliorate these issues, but limits themselves present certain problems, a detailed analysis of which is beyond the scope of this paper. Since the 1970s, a new approach to calculus, called non-standard analysis, has returned to the infinitesimal in a more rigorous form in an attempt to provide a more secure foundation to analysis (Robinson, 1974).

The problem presents itself in number theory as a deep conceptual gulf between the continuous and the discrete. At issue is the very different nature of the infinite conceived as composed of the natural numbers, versus that composed of the real numbers. As Cantor has proven, the real numbers are uncountable, and cannot be put into one-to-one correspondence with the infinite set of natural numbers (Cantor, 1874). Any attempt to enumerate the reals between say, 0 and 1, in terms of counting numbers, will fail, and not just because it would require infinite counting. Even if we assume an infinite series of natural numbers, the project will fail, because any attempt to count the reals will leave some out. Therefore, the continuous infinite both cannot be mapped onto, and is itself infinitely larger than, the discrete infinite.

The underlying issue leading to the inability to reconcile the continuous and the discrete is the same fundamental problem that occurs when trying to reconcile tensed and tenseless time. The natural numbers are the counting numbers, and counting is deeply interwoven with the past-present-future model of time, deriving from cognitive clocks that record intervals. We learn to count by keeping a record of the number of times we have registered examples of a set of objects, for example fingers. The real numbers derive

from measures of quantity, with regard to space, and duration, with regard to time, for example comparing the amounts of water in two glasses. When we attempt to map the naturals onto the reals the issues we run into derive directly from attempting to map a concept derived from a tensed conception of time onto one originating from a tenseless conception.

The continuous and discrete problem is expressed in physics in terms of the fundamental conflict between quantum mechanics and relativity (Hagar, 2014). Quantum mechanics admits only of discrete entities, or quanta, whether these be in terms mass, energy, charge, or even time. Relativity deals in continuous quantities for all of these, as did classical physics. Again, a detailed analysis of this issue cannot be entered into here, although it should be noted that the fact that the world appears continuous to us at large spatial and temporal scales, and discrete at small ones, is likely to be closely connected to our evolved mechanisms for integrating our perceptual interpretations of sense data into our cognitive models of space and time. These mechanisms were acquired as tools necessary for the full expression of theory of mind and episodic memory.

As mentioned above, the term “stuck in time” has been used in a general sense to describe the temporal experience of non-human animals (Roberts, 2002), and this experience has been referred to as that of a “continuous present.” The term “continuous present” is perhaps the most succinct linguistic expression of the temporal paradox that is possible. The discussion above treated the tensed time model as a legitimate option, but in reality tensed time can yield neither the continuous nor the discrete. We can say the present yields continuity by enabling smooth temporal flow, but it does not, for it is

conceived as a moment. We can say the present yields discrete quanta by enabling us to count intervals, but it does not, for it provides no structural context for these intervals to inhabit. The concept of tensed time is incoherent in itself (McTaggart, 1908), and the paradox changes form if we accept that one of the traditional choices is fundamentally structurally unsound. Tenseless time has a more secure foundation in terms of its conceptual structure. While it cannot fully defuse the conflict between the continuous and the discrete, the tenseless model at least avoids directly denying both, but we find restraining our inclination to reject it based on contradiction with what we see as our actual experience to be extremely difficult.

If, for the purposes of this paper, it is accepted as axiomatic that an organism's mode of time perception derives originally from its adaptive utility in an evolutionary context, some progress can be made toward the resolution of this conflict. Considering humans, our denial of the plausibility of tensed time derives from conceptual analysis, whereas our rejection of tenseless time is based on intuition and our phenomenal experience. This disparity has certain implications.

The tenseless time model, in addition to its rigorous theoretical grounding, is supported at the fundamental levels of physics by relativity (Putnam, 1967), which is itself supported by extensive empirical evidence. From the perspective of physics, it is by far the dominant model . But however much we accept this evidence, tenseless time is still not empirically supported from the perspective of phenomenal experience. This suggests that our phenomenal representation of tensed time is an overlay built upon a tenseless substrate. The dependency of our conceptual schema on developed

metacognition, itself tied to our linguistic capability, prevents us from being able to describe this substrate coherently, for whenever we attempt to our description unavoidably involves the projection of tensed conceptual apparatus. Our equivocation of “stuck in time” with “continuous present” exemplifies this. We see that neither of these two concepts is fully coherent, and additionally that, even if they were, the equivocation is erroneous, but we cannot describe why this is in any way that is both logically rigorous and intuitively satisfying. Progress can be made by an exploration of how the evolution of temporal representation parallels the development of episodic memory and theory of mind.

### **3.2 The Evolution of Temporal Representation**

Taking the sequence of REM → MR → PEM → MC, each transition requires the elaboration of the organism's temporal representation. In similarity to the interdependence between episodic memory and theory of mind, this requires the importation of cognitive structures from another domain, that of spatial representation. The conceptual account of how this transpired is supported by neurological evidence.

In humans and other animals, theory of mind and episodic memory are correlated with activation of very similar neural networks, which themselves are closely tied to an underlying system connected to spatial navigation (Rabin, Gilboa, Stuss, Mar, and Rosenbaum, 2010). The necessity of spatial representation in remembering the *where* of “what, when, where” is self-evident, and the correlation between spatial representation and theory of mind could be explained, in the light of the analysis above, purely in terms of this connection to episodic memory. This, however, does not address the conceptual

issues raised by the involvement of spatial cognitive systems in the expression of theory of mind, and even regarding episodic memory, such involvement goes beyond the minimal level of remembering *where*, and is deeply connected with the developing representation of time.

As explained above, in order for an organism to mind-read, they must be able to generate a spatiotemporal matrix in which to place the representation of the target of their mind-reading. This matrix, with regard to events in what we would call the past, is already present and active in an organism capable only of retrospective episodic memory, but does not at that point require any representation of movement in time. There is no requirement to postulate that an organism capable of retrospective episodic memory experience anything resembling temporal flow.

The addition of an agent requires that the mind-reader develop, at least on a first-order level, the ability to differentiate which of their own episodic memories include or relate to this agent, and to replay these events in sequence in order to predict the current state of the target of their mind-reading. This requires that a minimal faculty that models temporal progression, which can be replayed as needed, be added to their representation of time.

The most parsimonious way for the organism to add this faculty is to import their spatial navigational system into their representation of time. There is no need at this point for the model of time as movement to be accompanied by any experiential state of temporal flow. As established above, the agent conception enabling mind-reading need not be a self-concept, and the ability to synthesize the cognitions of space and time in this

way can still be entirely context specific, in addition to being restricted to what we term the past.

Some mind-readers develop prospective episodic memory. In order for this to occur, the spatiotemporal matrix, with the included ability to replay sequences of events involving agents, must be duplicated, and the organism must gain the ability to populate this duplicate matrix with simulated events and simulated agents. This organism differs from the simple mind-reader in that the agent model has become unmoored from its connection to another individual, and must operate in a conceptual space where simulated events can be modeled without reference to specific preexisting memories. This simulation matrix must also be carefully differentiated from the actual WWW record of the organism's memories, lest massive confusion ensue. These requirements necessitate additions to the organism's representation of time.

In order to maintain the distinction between the simulation space and the actual memory matrix, a bifurcation of time representation is required. There must be a representation referring to the simulation space, dissociated from one referring to the memory record, and the ability to replay sequences of events using representations of movement imported from spatial cognition must be present in both. This is the start of a representation of what we will later term *the future*, but there is again no need for the organism to possess such a concept at this point.

This representational system now contains sufficient conditions for the development of a nascent self-concept, as the organism increasingly identifies the agent operating within the simulation space used for planning as a specific instantiation of the

model agent used for mind reading, and begins to project this agent representation into the original episodic memory matrix, the contents of which now begin to become *their* memories. The general qualities of this self-concept are derived from observation of other individuals in the social group, and the specific autobiographical events conditioning the concept of the self as unique individual derive from the increasing ability of the agent to actively explore the episodic memory matrix. At this point, the organism is beginning to come unstuck from time. Corvids and apes, in addition to other groups discussed above in the context of the mirror test, have likely to varying degrees reached this point.

The final step is the full hooking up of this nascent self-concept with the capabilities of the organism as mind-reader. Metacognition originates from the development of the ability to observe the operation of the entire system described above, from the perspective of an outside observer who now has privileged access to the episodic memory and mind-reading capabilities of their target, which is the same system from which they arose. This allows second-order cognition to occur.

In terms of temporal representation, this results in the organism identifying the representational systems enabling episodic memory and mind-reading with actual features of temporal reality, and simultaneously the ability to range over and inhabit both the original and simulated memory matrices. This is the origin of the explicit concepts of past, present, and future, and the ability to engage in explicit mental time travel between points within them as though they are points in space. These representations, coupled with the experience of temporal flow modeled on spatial movement, immediately begin

to generate the perplexing temporal paradoxes discussed above, because the representation of temporal flow requires that the self stand outside of that flow, and directly conflicts with the representation of events in time as points in space.

### **3.3 Temporal Representation in Unusual Human Mental States**

This model provides a possible explanation for the rigidity of our apparently fallacious view of time. Our metacognitive ability, and thus our ability to describe experiences in language, is built on a substructure which requires that we experience time this way, or else lose our sense of self, at least in its familiar form. However, there is some evidence from certain more unusual human experiences of time that, either fortuitously or with the correct training, it is possible for at least some components of this substructure to enter the field of consciousness. The theory predicts that any phenomenal experience of these cognitive foundations would obviate the concurrent presence of the metacognitive self, or alternatively lead to extremely unusual temporal representations arising as a result of combining the components into unfamiliar recursive systems.

The simplest method of changing temporal representation is to loosen the identification the self with the stream of second-order cognition. This is the state often known as *mindfulness*, where the phenomenal self takes an observational perspective, involving a modification of phenomenal consciousness such that the second order flow is not experienced as constituting the self (Vago and Silbersweig, 2012). In the conceptual scheme described above, this can be theorized to involve phenomenal awareness of the recursive relationship between second-order cognitions and the self-concept.

Against immediate intuition, this should not be described as the addition of a

third-order cognitive stream. The mindful self does not form and identify with a new flow of cognitions. Rather, the continuous self-interpretation, or internal mind-reading, underlying the second-order stream begins to enter the domain of conscious attention. In terms of affective empathy, this is experienced as the development of an explicit affective understanding and compassion, rather than implicit mirroring, of the emotional states of the second-order self (Vago and Silbersweig, 2012).

Mindfulness must also not be misinterpreted as a return to the nascent self concept of the non-metacognizer, which is more likely shown in the experience of flow states, as explained above in the description of neanderthal cognition (Wynn and Coolidge 2004). In the case of mindfulness, the second-order stream remains in place, but second-order cognitions are no longer experienced as identical to the mental states of the self. In addition to many psycho-therapeutic benefits, this state begins to provide some intuition that temporal representation is more flexible than is usually apparent, as the mindful self begins to observe the metacognitive flow underlying the experience of the perpetually transforming present, in contrast to the full immersion of the second-order self in this flow (Kramer, Weger, and Sharma, 2012).

Advanced meditators often report that the field of conscious attention can expand further, to include additional awareness of the substructure of spatiotemporal representation underlying and enabling the second-order self's experience of temporal flow. This is also sometimes elicited by psychedelic drugs such as psilocybin when the subject is in the right state of mind and environmental context (Wittmann et al., 2007). Descriptions of this state as “timeless,” and the frequent identification of this with

“selfless,” provide support to the hypothesis that the temporal representation of an organism without metacognition is significantly different to the usual human experience. As in the previous example, this state must not be identified as a return to the phenomenal state of such an organism, it is an expansion rather than a diminution of the field of phenomenal awareness (Raffone and Srinivasan 2010).

Such states are also often described as giving insight into the fundamental unity of consciousness across multiple individuals, along with an enhancement of both cognitive and affective empathy (Raffone and Srinivasan 2010). This is evidenced in the use of terms such as “oneness,” and “the universal self,” and, according to the hypothesis presented in this paper, reflects the similarity of this state with that of the non-metacognitive mind-reader, who could be described as existing in a state of selfless affective and cognitive empathy. Additionally, the form of attachment of individuals to their own episodic memories is modified in such states of mind. Sometimes people even report telepathic experiences, which would correlate with this reduced sense of ownership of their memories, and awareness of the fundamental role of empathy in the construction of the self-concept (Raffone and Srinivasan 2010).

Accessing still more basic levels of representation is likely to only be possible with a certain degree of fragmentation and recursive recombination of the entire structure, and minds that explore these areas are unlikely to be able to report their experiences using language, but again, there is evidence that this also occurs among experienced meditators, and in certain psychedelic experiences, where variable experiences of standing outside of space and time sometimes occur in the midst of

extensive modification, disruption, or near absence of the agential self-concept, and the loss of the subject-object distinction (Grittiths et al., 2006). Alterations in temporal representation, episodic memory, and theory of mind, are also commonly seen in mental illness, for example in depersonalization and dissociative disorders (Simeon and Abugel, 2006), among others, and disturbance of theory of mind is a fundamental component of autism (Baron-Cohen, 1997). In all of these cases there are correlated changes in the experiential self-concept.

As a final observation, there is some evidence that, even remaining within the confines of syntactic tensed language, there is more flexibility in temporal representation than may be apparent. Although language and metacognition binds us to the representation of time as either a space within which we move, and within which events occur, or a continual and protean experiential flow, there are some languages that show we have more options over the specifics of the paradoxical binding of these representations than we might expect. In some languages, time is described as approaching from behind and receding in front of us (Karen and Bui, 2016). That is, the future comes up from behind, which, given that we can *see* the past in memory, but have no such access to the future, certainly makes sense conceptually. There is neither space nor time for a detailed analysis of this perplexing observation at this point. The reader is encouraged to remain as perplexed as the author will always be in their continual exploration of the hidden substructure conditioning our phenomenal reality.

## Conclusion

This paper has presented a specific conceptual model of the interdependent evolution of episodic memory and theory of mind, which additionally accounts for fundamental features of our temporal representation, and is amenable to empirical testing. The model makes several specific predictions, and is grounded in providing an account of the development of the human self-concept and experience of time that is evolutionarily plausible. This theory accounts for the strong evidence that the fundamental nature of time is something quite different from general human experience, and provides specific interpretations of the representational background to unusual human experiences of time and the self. The paper has the additional aim of showing the necessity of a detailed analysis of the evolution and representational structure of episodic memory in any account of theory of mind.

Further exploration of the precise pattern of linkage between episodic memory, theory of mind, and temporal representation, would serve to provide data to test the central arguments of this paper. Ideally this would include demonstrated correlation between phenomenological, behavioral, and neurological levels of analysis, and exploration of the specific neural substrates of the faculties, which would serve to demonstrate whether their structural connectivity concords with the hypotheses presented here.

## References

- Abramson, C. I., & Chicas-Mosier, A. M. (2016). Learning in plants: Lessons from *Mimosa pudica*. *Frontiers in Psychology*, 7(417). doi: 10.3389/fpsyg.2016.00417
- Ambrose, S. H. (2010) Coevolution of Composite-Tool Technology, Constructive Memory, and Language: Implications for the Evolution of Modern Human Behavior. *Current Anthropology* 51(1), 135-147.
- Ari, C. & D'Agostino, D. P. (2016). Contingency checking and self-directed behaviors in giant manta rays: Do elasmobranchs have self-awareness? *Journal of Ethology*, 34(2), 167–174.
- Azéma, M., & Rivière, F. (2012). Animation in Palaeolithic art: A pre-echo of cinema. *Antiquity*, 86(332), 316-324. doi: 10.1017/S0003598X00062785
- Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. Cambridge, Massachusetts: MIT Press.
- Bartal1, I. B-A., Decety, J., & Mason, P. (2011). Empathy and pro-social behavior in rats. *Science*, 334(6061), 1427-1430. doi: 10.1126/science.1210789
- Beyin, A. (2011). Upper Pleistocene human dispersals out of africa: A review of the current state of the debate. *International Journal of Evolutionary Biology*, 2011, 615094. doi: 10.4061/2011/615094
- Blumenfeld, R. S., & Ranganath C. (2007). Prefrontal cortex and long-term memory encoding: An integrative review of findings from neuropsychology and neuroimaging. *The Neuroscientist*, 13(3), 280-91. doi:

10.1177/1073858407299290

- Boisseau, R. P., Vogel, D., & Dussutour, A. (2016). Habituation in non-neural organisms: evidence from slime moulds. *Proceedings of the Royal Society B: Biological Sciences*, 283(1829). doi: 10.1098/rspb.2016.0446
- Brandimonte, M. A., Einstein, G. O., & McDaniel, M. A. (1996). *Prospective Memory: Theory and Applications*. New York: Psychology Press.
- Broly, P., & Deneubourg, J-L. (2015). Behavioural contagion explains group cohesion in a social crustacean. *PLoS Computational Biology*, 11(6). doi: 10.1371/journal.pcbi.1004290
- Burghardt, G. M. (2004). Iguana research: looking back and looking ahead. In A. C. Alberts, R. L. Carter, W. K. Hayes, & E. P. Martins (Eds.), *Iguanas: Biology and Conservation* (pp. 1-12). Berkeley: University of California Press.
- Cammaerts, M-C., & Cammaerts, R. (2015). Are ants (Hymenoptera, Formicidae) capable of self recognition? *Journal of Science*, 5(7), 521-532.
- Cantor, G. (1874). Ueber eine eigenschaft des inbegriffes aller reellen algebraischen zahlen. *Journal für die Reine und Angewandte Mathematik*, 77, 258–262.
- Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, 32(2), 121-138.
- Clayton, N. S., & Dickinson, A. (1998) Episodic-like memory during cache recovery by scrub jays. *Nature*, 395, 272–274.
- Clayton, N. S., Salwiczek, L. H., Dickinson, A. (2007). Episodic memory. *Current Biology*, 17(6), 189–191.

- Dally, J. M., Emery, N. J., & Clayton, N. S. (2010). Avian theory of mind and counter espionage by food-caching western scrub-jays (*Aphelocoma californica*). *European Journal of Developmental Psychology*, 7(1), 17-37.
- Edgar, J. L., Lowe, J. C., Paul, E. S., & Nicol, C. J. (2011). Avian maternal response to chick distress. *Proc. R. Soc. B.* doi: 10.1098/rspb.2010.2701
- Fernández-Laso, M.C., Navarro, M. G. C., García-Antón, M. D., & Rivals, F. (2011). Territorial mobility of Neanderthal groups: A case study from Level M of Abric Romaní (Capellades, Barcelona, Spain). In N. J. Conard, & J. Richter (Eds.) *Neanderthal Lifeways, Subsistence and Technology: One Hundred Fifty Years of Neanderthal Study*. Dordrecht: Springer.
- Finlayson, C., Brown, K., Blasco, R., Rosell, J., Negro, J. J., Bortolotti, G. R., Finlayson, G., Marco, A. S., Pacheco, F. G., Vidal, J. R., Carrión, J. S., Fa, D. A., & Llanes, J. M. R., (2012). Birds of a feather: Neanderthal exploitation of raptors and corvids. *PLoS ONE*, 7(9), 45927. doi: 10.1371/journal.pone.0045927
- Gallese, V., & Goldman, A. (1998). Mirror Neurons and the Simulation Theory of Mind-reading. *Trends in Cognitive Sciences*, 2(12), 493-501.
- Gallup, GG Jr. (1970). Chimpanzees: Self recognition. *Science*, 167(3914), 86–87.
- Gibbons, A. (2011) Who were the Denisovans? *Science*, 333(6046), 1084-1087. doi: 10.1126/science.333.6046.1084
- Graf, P. & Utzl, B. (2001) Prospective memory: a new focus for research. *Consciousness and Cognition*, 10(4), 437-450.
- Griffiths, R. R., Richards, W. A., McCann, U., Jesse, R. (2006). Psilocybin can occasion

- mystical-type experiences having substantial and sustained personal meaning and spiritual significance. *Psychopharmacology*, 187, 268. doi: 10.1007/s00213-006-0457-5
- Hagar, A. (2014). *Discrete or Continuous? The Quest for Fundamental Length in Modern Physics*. Cambridge: Cambridge University Press.
- Held, S., Mendl, M., Devereux, C., & Byrne, R.W. (2001). Behaviour of domestic pigs in a visual perspective taking task. *Behaviour*, 138(11), 1337–1354.
- Heyes, P. J., Anastasakis, K., de Jong, W., van Hoesel, A., Roebroeks, W., & Soressi, M. (2016). Selection and use of manganese dioxide by Neanderthals. *Nature Scientific Reports*, 6, 22159.
- Higham, T., Douka, K., Wood, R., Ramsey, C., Brock, F., Basell, L., Camps, M., Arrizabalaga, A., Baena, J., Barroso-Ruiz, C., Bergman, C., Boitard, C., Boscato, P., Caparrós, M., Conard, N., Dräily, C., Froment, A., Galván, B., Gambassini, P., García-Moreno, A., Grimaldi, S., Haesaerts, P., Holt, B., Iriarte-Chiapusso, M., Jelinek, A., Jordá Pardo, J., Maíllo-Fernández, J., Marom, A., Maroto, J., Menéndez, M., Metz, L., Morin, E., Moroni, A., Negrino, F., Panagopoulou, E., Peresani, M., Pirson, S., de la Rasilla, M., Riel-Salvatore, J., Ronchitelli, A., Santamaría, D., Semal, P., Slimak, L., Soler, J., Soler, N., Villaluenga, A., Pinhasi, R. and Jacobi, R. (2014). The timing and spatiotemporal patterning of Neanderthal disappearance. *Nature*, 512(7514), 306–309.
- Horowitz, A. (2008). Attention to attention in domestic dog (*Canis familiaris*) dyadic play. *Animal Cognition*, 12(1), 107–18.

- Jonides, J., Lewis, R. L., Nee D. E., Lustig, C. A., Berman, M. G., & Moore K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology, 59*, 193-224. doi: 10.1146/annurev.psych.59.103006.093615
- Jozet-Alves, C., Bertin, M., & Clayton, N. S. (2013). Evidence of episodic-like memory in cuttlefish. *Current Biology, 23*(23), 1033-1035.
- Kaminski, J., Call, J., & Tomasello, M. (2006). Goats' behaviour in a competitive food paradigm: Evidence for perspective taking? *Behaviour, 143*(11), 1341–1356.
- Kart-Teke, E., Silva, M. A. D. S., Huston J. P., & Dere E. (2006). Wistar rats show episodic-like memory for unique experiences. *Neurobiology of Learning and Memory, 85*(2), 173-182.
- Kirk, K. G. S., Raven, J. E., Schofield, M. (1983). *The Presocratic Philosophers. A Critical History with a Selection of Texts, Second Edition*. Cambridge: Cambridge University Press.
- Kouwenberg, A-L., Walsh C. J., Morgan B. E., & Martin G. M. (2009). Episodic-like memory in crossbred Yucatan minipigs (*Sus scrofa*). *Applied Animal Behaviour Science, 117*(3), 165-172.
- Kozowyk, P. R. B., Soressi, M., Pomstra, D., Langejans G. H. J. (2017). Experimental methods for the Palaeolithic dry distillation of birch bark: implications for the origin and development of Neandertal adhesive technology. *Scientific Reports, 7*(1) doi: 10.1038/s41598-017-08106-7
- Krachun, C., Carpenter, M., Call, J., & Tomasello, M. (2009). A competitive nonverbal false belief task for children and apes. *Developmental Science, 12*(4), 521-535.

- Kramer, R. S., Weger, U., W., & Sharmaa, D. (2013). The effect of mindfulness meditation on time perception. *Consciousness and Cognition*, 22(3), 846-852.
- Maginnity, M. E., & Grace, R. C. (2014). Visual perspective taking by dogs (*Canis familiaris*) in a Guesser–Knower task: evidence for a canine theory of mind? *Animal Cognition*, 17(6), 1375–1392.
- Martin-Ordas, G., Haun, D., Colmenares, F. & Call, J. (2010). Keeping track of time: Evidence for episodic-like memory in great apes. *Animal Cognition*, 13(2), 331-340.
- Mathis, R., & Ackermann, M. (2016). Response of single bacterial cells to stress gives rise to complex history dependence at the population level. *Proceedings of the National Academy of Sciences of the United States of America*, 113(15), 4224-9.  
doi: 10.1073/pnas.1511509113
- McHugh, L. & Stewart, I. (2012). *The Self and Perspective-taking: Contributions and Applications from Modern Behavioral Science*. Oakland, CA: New Harbinger.
- McTaggart, J. M. E. (1908). The Unreality of Time. *Mind*, 17, 457–73.
- Mitchell, P. (2011). Acquiring a theory of mind. In A. Slater & J. Gavin Bremner (Eds.), *An Introduction to Developmental Psychology, 3rd ed.* (pp. 381–406). Hoboken, New Jersey: John Wiley & Sons Inc.
- Montemayor, C. (2013). *Minding Time: A Philosophical and Theoretical Approach to the Psychology of Time*. Boston: Brill.

- Morgan, G., & Kegl., J. (2006). Nicaraguan sign language and theory of mind: The issue of critical periods and abilities. *Journal of Child Psychology and Psychiatry*, 47(8), 811-819.
- Pahl, M., Zhu, H., Pix, W., Tautz, J., & and Zhang, S. (2007). Circadian timed episodic-like memory – A bee knows what to do when, and also where. *Journal of Experimental Biology*, 210(20), 3559-3567.
- Pastukhov, A. (2016). Perception and the strongest sensory memory trace of multi-stable displays both form shortly after the stimulus onset. *Attention, Perception, & Psychophysics*, 78(2), 674-84. doi: 10.3758/s13414-015-1004-4.
- Perry, A., & Shamay-Tsoory, S. (2013). Understanding emotional and cognitive empathy: A neuropsychological perspective. In S. Baron-Cohen, H. Tager-Flusberg, & M. Lombardo (Eds.), *Understanding Other Minds: Perspectives from Developmental Social Neuroscience* (p.178-208). Oxford: Oxford University Press.
- Petri H. L., & Mishkin M. (1994). Behaviorism, cognitivism and the neuropsychology of memory. *American Scientist*, 82(1), 30-37.
- Plotnik, J.M., de Waal, F.B.M., & Reiss, D. (2006). Self-recognition in an Asian elephant. *Proceedings of the National Academy of Sciences of the United States of America*, 103(45), 17053–17057.
- Povinelli, D. J., Rulf, A. B., Landau, K. R. & Bierschwale, D.T. (1993). Self-recognition in chimpanzees (*Pan troglodytes*): distribution, ontogeny, and patterns of emergence. *Journal of Comparative Psychology*, 107(4), 347–372.

- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, special issue: Cognition and Consciousness in Nonhuman Species*, 1(4), 515–526.
- Prior, H., Schwarz, A. & Güntürkün, O. (2008). Mirror-induced behavior in the magpie (*Pica pica*): Evidence of self-recognition. *PLoS Biology*, 6(8), 202.
- Putnam, H. (1967). Time and Physical Geometry. *Journal of Philosophy*, 64, 240–247.
- Rabin, J. S., Braverman, A., Gilboa, A., Stuss, D. T., & Rosenbaum, R., S. (2012). Theory of mind development can withstand compromised episodic memory development. *Neuropsychologia*, 50(14), 3781-3785.
- Raffone, A. & Srinivasan, N. (2010). The exploration of meditation in the neuroscience of attention and consciousness. *Cognitive Processing*, 11(1). doi: 10.1007/s10339-009-0354-z
- Reiss, D. & Marino, L. (2001). Mirror self-recognition in the bottlenose dolphin: A case of cognitive convergence. *Proceedings of the National Academy of Sciences of the United States of America*, 98(10), 5937–5942
- Rendu W., Beauvalc, C., Crevecoeur, I., Bayled, P., Balzeau, A., Bismuthb, T., Bourguignong, L., Delfourd, G., Faivred, J-P., Lacrampe-Cuyaubère, F., Tavorminac, C., Todiscoj, D., Turqd, A., & Maureilled, B. (2014). Evidence supporting an intentional Neandertal burial at La Chapelle-aux-Saints. *Proceedings of the National Academy of Sciences of the United States of America*, 111(1), 81-86.

- Roberts, W. W. (2002). Are animals stuck in time? *Psychological Bulletin*, 128(3), 473-489.
- Robinson, A. (1974). *Non-standard Analysis*. Princeton: Princeton University Press.
- Roebroeksa, W., Siera, M. J., Nielsena, T. K., De Loeckera, D., Parésb, J. M., Arpsd C. E. S., & Müchere, H. J. (2012). Use of red ochre by early Neandertals, *Proceedings of the National Academy of Sciences of the United States of America*, 109(6), 1889-1894.
- Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, 392, 595-598.
- Sadiera, B., Delannoya, J-J., Benedettib, L., Bourlèsb, D. L., Jailletab, S., Genestec, J-M., Lebatardb, A-E., & Arnoldb, M. (2012). Further constraints on the Chauvet cave artwork elaboration. *Proceedings of the National Academy of Sciences of the United States of America*, 109(21), 8002-8006.
- Salwiczek, L. H, Watanabe, A. & Clayton., N. S. (2010). Ten years of research into avian models of episodic-like memory and its implications for developmental and comparative cognition. *Behavioural Brain Research*, 215(2), 221-234.
- Sato, N., Tan, L., Tate, K., & Okada M. (2015). Rats demonstrate helping behavior toward a soaked conspecific. *Animal Cognition*, 18(1039). doi: 10.1007/s10071-015-0872-2

- Simeon, D., & Abugel, J. (2006). *Feeling Unreal: Depersonalization Disorder and the Loss of the Self*. Oxford: Oxford University Press.
- Stapert, D. (2007). Neanderthal children and their flints. *Journal of Archaeology of Northwest Europe*, 1(2).
- Sullivan, K. & Bui. L. T. (2016). With the future coming up behind them: Evidence that time approaches from behind in Vietnamese. *Cognitive Linguistics*, 27(2), 205-233.
- Terrace, H. S., (2005). Metacognition and the origin of language. In H. S. Terrace, & J. Metcalfe (Eds.), *The Missing Link in Cognition: Origins of Self-Reflective Consciousness* (p. 84). Oxford: Oxford University Press.
- Tulving, E. (1985). *Elements of Episodic Memory*. Oxford: Oxford University Press.
- Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology*, 53, 1-25.
- Uchino E, Watanabe S. (2014). Self-recognition in pigeons revisited. *Journal of Experimental Animal Behavior*, 102(3), 327-34. doi: 10.1002/jeab.112
- Vago, D. R. & Silbersweig, D. A. (2012). Self-awareness, self-regulation, and self-transcendence (S-ART): A framework for understanding the neurobiological mechanisms of mindfulness. *Frontiers in Human Neuroscience*, 6, 296. doi: 10.3389/fnhum.2012.00296
- Washburn, D. A., Gulledge, J. P., Beran, M. J., & Smith, D. (2010). With his memory

- magnetically erased, a monkey knows he is uncertain. *Proceedings of the Royal Society: Biology Letters*, 6, 160-162. doi: 10.1098/rsbl.2009.0737
- Westergaard, G. C., & Hyatt, C. W. (1994). The responses of bonobos (*Pan paniscus*) to their mirror images: Evidence of self-recognition. *Human Evolution*, 9.
- Whiten, A. (2001). Theory of mind in non-verbal apes: Conceptual issues and the critical experiments. *Philosophy*, 199.
- Wimmer, H., Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103–128.
- Wittmann, M., Carter, O., Hasler, F. B., Cahn, R., Grimberg, U., Spring, P., Hell, D., Flohr, H., & Vollenweider, F. X. (2007). Effects of psilocybin on time perception and temporal control of behaviour in humans. *Journal of Psychopharmacology*, 21(1), 50-64
- Wynn, T., & Coolidge, F. L. (2004). The expert Neandertal mind. *Journal of Human Evolution*, 46(4), 467-487.
- Zentall, T. R., Singer, R. A., & Stagner, J. P. (2008). Episodic-like memory: Pigeons can report location pecked when unexpectedly asked. *Behavioural Processes*, 79(2), 93-98.
- Zilhão, J. (2007). The emergence of ornaments and art: An archaeological perspective on the origins of “behavioral modernity.” *Journal of Archaeological Research*, 15(1), 1–54.