Phylogeny and ontogeny of mental time
Satoshi Hirataa, Toru Betsuyakub, Kazuo Fujitac, Tamami Nakanoc, Yuji Ikegayad,e

a Wildlife Research Center, Kyoto University, Kyoto 606-8203, Japan
b Graduate School of Letters, Kyoto University, Kyoto 606-8501, Japan
c Graduate School of Frontier Biosciences, Osaka University, Suita, Osaka, 565-0871, Japan
d Graduate School of Pharmaceutical Sciences, The University of Tokyo, Tokyo 113-0033, Japan
e Center for Information and Neural Networks, National Institute of Information and Communications Technology, Suita City, Osaka, 565-0871, Japan

Keywords: Mental time, Episodic memory, Episodic foresight, Phylogeny, Ontogeny, Mental time travel

Abstract
Humans have mental time in our mind, apart from physical time that is a part of system that governs the physical world, and memory is our key cognitive ability for recognizing the passage of time. Recent studies have suggested that the memory system of several nonhuman animals may have an incidental nature, which is also a feature of episodic memory. In addition, apes, which are phylogenetically close to humans, have an ability to remember a single past event. In the case of humans, preverbal infants under the age of two are able to retain long-term memory of a single event and apply it to predict a future event. Thus, nonhuman animals and preverbal human infants both have their own specific mental time travel abilities, and there is a phylogenetic and ontogenetic basis of full-fledged mental time travel that can be found in human adults.

1. General introduction of the mental chronogenesis

Why do we have memories of past events, but not of future events? This question might seem like a superficial word game, but it is actually rooted in the more essential question about the “arrow of time,” i.e., “why does time flow in one direction from the past to the future?” Moreover, it is likely related to the philosophical question: “why does time exist?” Whatever the reason, time exists in this world, and our brain has evolved to be able to sense time.

Consider, for a moment, a world that does not change at all. The hands of all clocks stop, no single leaf sways and all surrounding objects are stationary. How would someone living there be able to tell the passage of time? If you were living in a world where nothing is changing ever since you were born, perhaps you would not feel (or understand) time. Thus, one could further ask: “why can you feel that time is passing?” The key to this question is memory.

We can only know the passing of time through the observation of things that change. Memory is the necessary function that allows us to perceive change. In order to detect the difference between past and present (i.e., change), we must remember how the world was like before. Only when you have your memory preserved, can you compare the past memory and the present situation to judge whether something has changed, by finding some difference between the two. If this is on track, we feel the passage of time because our brain has memory. Consistent with this idea, patients with hippocampal lesions do not only exhibit anterograde amnesia but also often exhibit a loss of the sense of time (Scoville and Milner, 1957)

Memory has the function of accumulating what we have seen in the past. The brain can store memories only because time flows from the past to the future in one direction. This is indeed a part of the definition of memories. If time were to flow backwards, memories in the brain would disappear in succession.

These thought experiments are puzzling, awkward and quite fun. However, they are still problematic because the premise is incomplete. How does one guarantee that time does not flow back-wards? In fact, there seems to be no physical necessity to assume that time flows in one direction, since according to modern physics, time may flow in both directions towards the future or the past (Price, 2012; Leifer and Pusey, 2017; Lesovik et al., 2019).

To recapitulate, we can sense “the change of time” because we have memories, since without memories we cannot observe change. At the same time, because memory integrates information through time, it is gradually removed if time were to pass in the opposite direction. In other words, if there were a backward flowing time it would not be sensed by our brains. As long as the brain detects time through memory, the only time that humans could observe is the time that flows from the past to the future.

As in the opening question, if we had a “memory of the future,” we would have had a rich sense of time that flows in both directions. The aim of this paper is to review a variety of experimental data and hypothesis about mental times from a comparative perspective.

2. Retrieval of incidentally encoded memory in non-primate species

Episodic memory is considered to be a crucial part of mental time travel, which is one of the noteworthy features of the human mind. Suddendorf and Corballis (1997) hypothesized that humans are capable of mental travel both to the past (episodic memory) and to the future (foresight), whereas nonhuman animals do not have this capability. Tulving (1985,2005) also made a similar claim. The question as to whether mental time travel is uniquely human has fueled comparative research in nonhuman animals.

Among the various behavioral paradigms used to assess episodic memory in nonhumans, the most representative is the WWW memory task (e.g. Clayton and Dickinson, 1998), based on the original notion of Tulving (1972) in which temporally dated episodes or events and temporal-spatial relations among these events are stored in episodic memory. It involves questioning whether animals can retrieve integrated memories that include information on content (What), location (Where), and temporal context (When). Although Tulving (1985) later revised the definition by acknowledging that the WWW criterion is neither necessary nor sufficient to define episodic memory, it is still seen as an important component of episodic memory in humans, thus researchers have applied the WWW criterion to test nonhuman animals as well. Through this approach, there are many successful demonstrations in various species (see Clayton, 2017; Dere et al., 2019; Crystal, 2018 for recent reviews of episodic memory in nonhumans).
To distinguish whether past events have been encoded explicitly or incidentally, however, the WWW approach is not adequate. If animals encode past events explicitly, it is possible they solve the task by non-episodic or semantic rules, by predicting the upcoming test. Another approach ensures that incidental encoding is required in order to clearly rule out a semantic account. It involves investigating whether animals spontaneously retrieve the past when unexpectedly asked (e.g. Zentall, 2013).

There is some evidence, albeit scarce, that this latter approach is feasible. For example, language-trained dolphins were able to repeat their recent actions when unexpectedly asked to do so (Mercado et al., 1998). In an operant procedure, pigeons successfully answered to an “unexpected question” test on an event recently encoded in another task (Zentall et al., 2001, 2008; Singer and Zentall, 2007). In a radial arm maze, rats chose the correct arm in an unexpected test that combined two tasks on which they were previously trained separately (Zhou et al., 2012). After a single experience, male meadow voles (Microtus pennsylvanicus) revisited the location where they previously found a pregnant female who would go into postpartum estrus in 24 h, depending on the passage of time since they last saw her (Ferkin et al., 2008). In a recent report using a deferred imitation task, dogs successfully recalled human actions even when asked unexpectedly (Fugazza et al., 2016), although the past episode used was not their own experience.

All the examples described above show that animals could retrieve their memory when asked in an unexpected manner, satisfying the incidental properties of encoding by ruling out the use of semantic rules. However, these previous studies have several problems that make it difficult to compare across species. For example, some studies combine several tasks which require intensive training. Others use highly species-typical behaviors of the subject species. To adequately evaluate diversity and the possible evolutionary history of episodic memory, comparisons across species appear essential. Therefore, a less-demanding task using species-general behavior, one sufficiently easy to enable comparisons among multiple species would be valuable.

In a study of dogs, Fujita et al. (2012) proposed a behavioral task that may satisfy the aforementioned requirement. In their first experiment, dogs, led by their owners, visited four baited containers in sequence, eating from two of them (exposure phase). At the remaining containers their owners prevented them from eating. After a delay of more than 10 min, dogs were unexpectedly released and allowed to freely re-explore the containers, now all empty (test phase). Because each dog experienced the foraging episode only once during the exposure phase, there was no chance of anticipation of being tested at a later point. Contrary to the non-episodic prediction (i.e. associative learning predicts that dogs would revisit where they had previously been fed), most dogs first revisited a container where they had not eaten, suggesting that they spontaneously recalled and used the incidentally formed memory of their past experience. A subsequent experiment further revealed that incidental memory in dogs contains information both on content (“what”) and location (“where”).

The experimental procedure in Fujita et al. (2012) may have multiple advantages for applying the “unexpected question” paradigm across species. The following paragraphs describe results from studies in our lab which have used Fujita’s et al. (2012) approach with several species, including cats and degus (Octodon degus).

In the case of cats (Takagi et al., 2017), we adopted almost the same procedure as the one used with dogs, except that instead of using a leash, the owners controlled their cat’s behavior by direct handling. Results showed that, like dogs, cats preferentially explored the feeder they had not eaten from, suggesting that they also used the incidentally encoded memory of their (non-)feeding experience. Additional experiments also revealed that this memory contains both “what” and “where” information.

For degus (Betsuyaku and Fujita, unpublished data), a four-armed radial maze with a feeder at the end of each arm was used. To control their foraging experience, instead of having direct contact with humans as in the cases of dogs and cats, some feeders were fit- ted with a false bottom to prevent foraging. Two of the four feeders were baited, but the bait in one was uncollectable because of the false bottom. Subjects visited all feeders once and were allowed to eat only from the baited-collectable feeder (exposure phase). After 5 min, they re-explored the apparatus with the remaining food removed (test phase). However, unlike dogs and cats, degus showed no preferential revisits or exploration of a feeder where they had not eaten.

In an alternative, non-foraging task of the WWW memory paradigm, degus were capable of recollecting the integrated memory of content (“what”) and location (“where”), along with the environmental context (“on which occasion”) (Betsuyaku et al., 2016). This finding, taken together with other positive reports in rodents (i.e. Zhou et al., 2012; Ferkin et al., 2008), suggests that some methodological issues might need to be dealt with before reaching a conclusion about incidental memory in degus. A crucial procedural aspect of our paradigm might be how to interfere with foraging by the animals in the exposure phase. Because dogs and cats are highly familiar with being controlled by humans, our tasks, as conceived, may be especially suitable for companion animals. But to expand the application to a wider range of species, appropriate procedures for behavioral control should be considered, with regard to the subject species’ natural repertoire.

One may argue that the behavior of dogs and cats might be an outcome of operant conditioning, because they are companion animals. Clearly, however, the functions of human episodic memory are not limited to feeding contexts.

3. Memories of a distant past and comprehension of the immediate past in apes

As indicated in the previous section, Fujita et al. (2012) suggested that, in addition to WWW characteristics, the incidental nature of events is an important aspect of episodic memories in humans. The aforementioned comparative research used foraging behavior to test for specific types of memories in nonhuman animals, and the WWW memory tests employed procedures which include some degree of training in which food was used as the reinforcer. Clearly, however, the functions of human episodic memory are not limited to feeding contexts.

Kano and Hirata (2015) developed a novel eye-tracking procedure to study incidental memory in chimpanzees and bonobos, the closest living great ape relatives of humans. They created a film for the apes and recorded eye movements as they watched it. In the first film, two actors appear on the scene and find bananas on the floor. Suddenly, another actor wearing a King-Kong suit appears from one of two doors, attacks one of the actors, steals the bananas, and escapes through the door. The directions of the apes’ eye movements were recorded, and after 24 h the apes watched the same film again. The prediction was that the apes would look at the door before the King Kong actor appeared, based on their memory of the film they had previously watched and on their anticipation of this critical event when watching it for the second time. The results confirmed this prediction.

In the second film produced by Kano and Hirata (2015), an actress and the King-Kong interact first, then the King-Kong attacks the actress. The actress had two potential weapons in front of her and used one of them to counterattack King-Kong. When the apes watched the film again within a 24-h delay, they focused on the tar-
weapon, even though this time the location of the two weapons was switched before the actress attempted to grab one. In sum, the apes made anticipatory visual scans based on what WHERE information, thus encoding ongoing events into long-term memory through a single experience. Suddendorf and Corballis (2010) suggest that the adaptive value of memories is to predict the future. The findings of Kano and Hirata (2015) suggest that apes also use their memory to predict future events.

An additional notable aspect of the study by Hirata and Kano (2015) is the complexity of the visual scene used as a video stim ulus, including changes over time. Most other work on episodic or episodic-like memory with non-humans uses quite limited events, and to the best of our knowledge one study involved a complex of a flow of events; Panoz-Brown et al. (2018) showed replay of memory for olfactory information in rats with up to 12 odors presented sequentially over several minutes.

Another related hypothesis is that in human development, episodic memory and recognition of one’s past actions are linked (Povinelli, 1995). Human infants typically begin to self-recognize in mirrors at 18–24 months of age, but for the same to occur in delayed video images, another two years are necessary (Amsterdam, 1972; Povinelli et al., 1996). Thus, children begin to understand their delayed self-image around the age of four, coinciding with their initial manifestations of autobiographical memory. Chimpanzees and other great apes have been shown to exhibit mirror self-recognition (Gallup, 1970), but there is little evidence regarding whether nonhumans could recognize themselves in images from different time periods.

Hirata et al. (2017) tested five chimpanzees with stickers placed on their head or face. Afterwards, either one of several self-videos including live and 1–4 s delayed videos was presented on a monitor, or one of the control videos was shown, each of which lacked a link to the subject’s current state (i.e. by showing non-self-images or self-images without stickers). The delay range of 1–4 s was chosen based on a study of human children which reported that 3-year-old children, during the transition from live self-recognition to the capacity for delayed self-recognition, do not remove stickers from their heads while viewing 2-s-delayed images of themselves, but do so while watching live or 1-s-delayed self-images (Miyazaki and Hiraki, 2006). Three of the five chimpanzees in Hirata et al. (2017) consistently removed the stickers while watching both live and 1–4 s delayed self-images on the monitor, but did so less frequently in control conditions. These findings suggest that the ability to extend one’s self-concept, in which one’s sense of identity extends across time and space, is not uniquely human, and may have evolved before the split of the chimpanzee-human lineage.

4. Development of episodic-like memory in human infants

In humans, the ability of mental time travel, composed of episodic memory and episodic foresight, becomes apparent in children around three to four years old, who can verbally recall the what-where-when of past events (Hayne and Imuta, 2011). Moreover, through episodic foresight, they can select an option in advance by anticipating future events based on past experiences (Atance and Sommerville, 2014; Suddendorf et al., 2011). However, mental time travel development in preverbal infants remains unclear because of the difficulty in examining it without the use of language (Clayton and Dickinson, 1998; Mulcahy and Call, 2006). It would not be surprising if it were the case that preverbal infants develop the precursors of episodic memory. Yet determining when and how such ability develops is an open question.

Various behavioral paradigms not involving verbal expressions have been proposed to investigate long-term episodic memory development in preverbal infants. The most representative paradigm is deferred imitation, which examines whether infants reproduce previously observed actions toward particular objects after delay. This paradigm revealed that infants as young as six months old can reproduce a particular action after a delay of 24 h. However, children younger than two could not reproduce the tar-get action unless it was repeatedly presented to them. Thus, it is an open question whether this ability of deferred imitation acts as a precursor of episodic memory, which is based on a single past event.

To examine the long-term memory of a single event in pre-verbal infants, Russell and Thompson (2003) applied a paradigm that demonstrated that scrub jays have an episodic-like memory of where and when a particular meal (“what”) was stored (Clayton and Dickinson, 1998). In their study, an experimenter placed a toy in each of three boxes while infants observed. Subsequently, the second experimenter removed a toy from one of the three boxes. After a 20 min and a 24-h delay, infants were encouraged to find a toy. Infants around two years old successfully selected the box in which the toy remained on the next day, but infants under 21 months did not. These results suggest that long-term event-based memory develops around the age of two. However, a problem with this task is that other cognitive abilities, in addition to event memory, are necessary for success. For example, infants have to hold a representation of object permanence across 24 h.

Further results of the Russell and Thompson (2003) experiment showed that 18-month-old infants preferentially selected the box from which the toy was removed after a delay of 20 min. This indicates that the observation of the toy-removal made infants associate the box with the availability of toys. The unmatured ability of suppressing such association might be the cause of poor performance in younger infants. Thus, it is possible that the development of episodic-like memory is underestimated in this task.

As described above, Kano and Hirata (2015) proposed an eye-tracking method to investigate long-term memory of one-time events in great apes. Since infants over three months can actively move their eyes to scan the external world, this method can be applied to various ages in preverbal infants, and it could clarify the developmental trajectory of anticipatory looks based on long-term event memory. Nakano and Kitazawa (2017) therefore presented the same film from the previous study of the great apes to 6-, 12-, 18-, and 24-month-old infants twice with an interval of 24 h and investigated whether they would display anticipatory examination of locations related to future events in their second viewing. As a result, 18- and 24-month-old infants increased their viewing time of the target location before the onset of the salient event in their second viewing. In contrast, 6- and 12-month-old infants did not change their looking behaviors between the first and second viewings. This suggests that by the age of 18 months, infants have developed long-term memory of “where” information of a one-time event.

In another scenario of the same experiment, 12-, 18-, and 24-month-old infants viewed a different video story, in which a human actor grabbed one of two tools to hit back at the ape character. In their second viewing after a 24-h delay, 18- and 24-month-old infants, but not 12-month-old infants, increased their viewing time of the tools before the human grabbed one. While 24-month-old infants preferentially looked at the tool that the human had used, 18-month-old infants did not show any preference for a particular tool. This indicates that 24-month-old infants formed a more precise memory of object information than did 18-month-old infants. Taken together, these results suggest infants at 18 months have developed long-term memory of a one-time event and this ability is elaborated over the following six months.

It is worth noting that there is a difference in anticipatory looking behavior between human infants and great apes although they viewed the same video stimuli in the same paradigm. Kano and Hirata (2015) reported that great apes made anticipatory looks just before the event onset when the salient cue was presented. In contrast, human infants started to show anticipatory looking behavior before the salient cue appeared. This difference implies that human infants form long-term memory of a relatively long series of events as a single episode compared to great apes.

5. Conclusion and future directions
Comparative research on nonhuman animals illustrates the question as to whether mental time travel into the past, namely episodic memory, is unique to human is complex. As shown by studies applying the unexpected question approach, the memory system of several species may have an incidental nature that is a feature of episodic memory. The studies on memories of single events using eye-tracking technology and delayed self-recognition also suggest that there is a continuity between human and non-human animals to a certain degree. When we look at humans, the notion that the ability of mental time travel—episodic memory and episodic foresight—emerges in children around three to four years of age has been challenged. Preverbal infants around 18 months can retain the long-term memory of a single event across 24 h and use it to forecast a future event. Moreover, infants around two years old manifest more accurate and elaborate event memory. Thus, the ability for mental time travel might have been formed during the evolutionary course by accumulating several subcomponents and a similar developmental trajectory might also occur in human ontogeny. Nonhuman animals and preverbal human infants both have specific mental time, and there is a phylogenetic and ontogenetic basis of full-fledged mental time which can be seen in human adults.

Cheng et al. (2016) proposed a new model regarding the evolutionary origin of mental time travel and suggested that scenario construction is an important component which only humans possess. A study by Saito et al. (2014) also suggested that there is a remarkable difference between humans and chimpanzees in their ability to imagine what they cannot see. Further comparative studies are needed to clarify the evolutionary origins of the ability of recalling the past and predicting the future.

Regarding the “unexpected question” approach, there are certainly remaining questions to be answered. For example, how long this specific type of memory lasts compared to other kinds of episodic memories is not fully understood. Because human episodic memory is often classified as part of long-term memory, the retention length of incidentally encoded memory should be revealed by means of experiments with proper a control to verify if this classification holds for this case. To evaluate the diversity of this kind of episodic memory among others, an important direction for future research is to compare similar tasks among multiple species including other mammals, birds (e.g. corvids), reptiles, fish, invertebrates as well as primates. Although the procedure might need modifications in response to each species’ peculiarities, managing to maintain the essential point of the task in Fujita et al. (2012) should be useful for our understanding of nonhuman research of episodic memories.

Finally, earlier anticipatory behavior in human infants compared to great apes, shown by Nakano and Kitazawa (2017), raises the possibility that systems of episodic-like memory are some—what different between preverbal human infants and nonverbal animals. It is possible that the interaction of episodic memory pre—cursors with the development of other abilities, such as language, object permanence and self-recognition, might facilitate the ability of mental time travel to develop into a more refined level during early childhood in humans.

Acknowledgments

This work was supported by JSPS Grants-in-Aid for Scientific Research (18H05520, 18H05522, 18H05524, 18H05525).

References