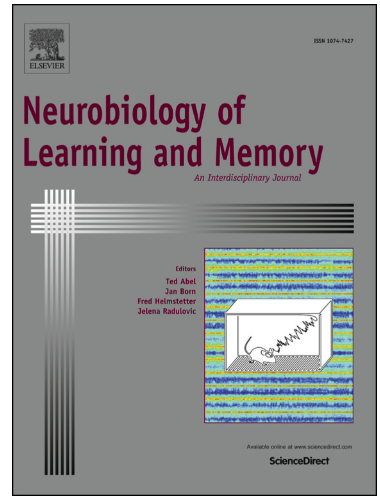
Accepted Manuscript



Disambiguating past events: Accurate source memory for time and context depends on different retrieval processes

Bjorn M. Persson, James A. Ainge, Akira R. O'Connor

PII: S1074-7427(16)30056-9
DOI: http://dx.doi.org/10.1016/j.nlm.2016.05.002
Reference: YNLME 6439

To appear in: *Neurobiology of Learning and Memory*

Received Date: 15 December 2015
Revised Date: 6 April 2016
Accepted Date: 8 May 2016

Please cite this article as: Persson, B.M., Ainge, J.A., O'Connor, A.R., Disambiguating past events: Accurate source memory for time and context depends on different retrieval processes, *Neurobiology of Learning and Memory* (2016), doi: http://dx.doi.org/10.1016/j.nlm.2016.05.002

TITLE: Disambiguating past events: Accurate source memory for time and context depends on different retrieval processes

RUNNING HEAD: Disambiguating the past using time and context

AUTHORS:

(1) Bjorn M. Persson [a]

(2) James A. Ainge [a]

(3) Akira R. O'Connor [a]

[a] School of Psychology & Neuroscience, University of St Andrews

St Mary's Quad, South Street

St Andrews, Fife

KY16 9JP

Scotland, United Kingdom

CORRESPONDING AUTHOR: James Ainge, School of Psychology and Neuroscience, University of St Andrews, St Mary's College, South Street, St Andrews, Fife, KY16 9JP, Scotland, UK. email: jaa7@st-andrews.ac.uk. Telephone: +44(0) 1334 462057

Abstract

Current animal models of episodic memory are usually based on demonstrating integrated memory for what happened, where it happened, and when an event took place. These models aim to capture the testable features of the definition of human episodic memory which stresses the temporal component of the memory as a unique piece of source information that allows us to disambiguate one memory from another. Recently though, it has been suggested that a more accurate model of human episodic memory would include contextual rather than temporal source information, as humans' memory for time is relatively poor. Here, two experiments were carried out investigating human memory for temporal and contextual source information, along with the underlying dual process retrieval processes, using an immersive virtual environment paired with a 'Remember-Know' memory task. Experiment 1 (n = 28) showed that contextual information could only be retrieved accurately using *recollection*, while temporal information could be retrieved using either *recollection* or *familiarity*. Experiment 2 (n = 24), which used a more difficult task, resulting in reduced item recognition rates and therefore less potential for contamination by ceiling effects, replicated the pattern of results from Experiment 1. Dual process theory predicts that it should only be possible to retrieve source context from an event using recollection, and our results are consistent with this prediction. That temporal information can be retrieved using familiarity alone suggests that it may be incorrect to view temporal context as analogous to other typically used source contexts. This latter finding supports the alternative proposal that time since presentation may simply be reflected in the strength of memory trace at retrieval – a measure ideally suited to trace strength interrogation using familiarity, as is typically conceptualised within the dual process framework.

Keywords: Episodic memory; time; context; recollection; familiarity

1. Introduction

The episodic memory system has been proposed to underpin our abilities to retrieve temporal-spatial relations between events, the subjective experience of reliving an event, autonoetic consciousness (Tulving, 1983), and mental time travel (Suddendorf & Busby, 2003). As conscious recollection and mental time travel are difficult to demonstrate in non-human animals, some researchers have suggested that episodic memory is a uniquely human cognitive ability (Suddendorf & Busby, 2003; Tulving, 2002). The utility of animal models is clear, as they allow for the study of episodic memory on a cellular and neural level currently not attainable in humans. Consequently a broad body of research has focused on the content, rather than the experience, of episodic memory to develop animal models (Clayton & Dickinson, 1998; Eacott & Easton, 2012) with almost exclusive emphasis placed on the integration of what happened, where it happened, and when (Clayton & Dickinson, 1998) or on which occasion it happened (Eacott & Easton, 2010; Eacott & Norman, 2004). According to these criteria, episodic memory and has been reported in rats (Eacott & Norman, 2004; Langston & Wood, 2010; Ergorul & Eichenbaum, 2004; Babb & Crystal, 2005, 2006; Kart-Teke et al. 2006), mice (Davis et al. 2013), birds (Clayton & Dickinson, 1998), chimpanzees (Martin-Ordas et al. 2013) and cuttlefish (Jozet-Alves et al. 2013). In the absence of a demonstration of subjective experience to definitively show correspondence across animal and human memory systems, the memory capability demonstrated in the what-where-when (WWWhen) and what-where-which (WWWhich) memory paradigms has been termed episodic-like memory (Clayton & Dickinson, 1998).

The use of these episodic-like memory paradigms has led to further debate as to whether events are separated in memory using solely temporal information, as suggested by WWWhen memory (Babb & Crystal, 2005, 2006; Clayton & Dickinson, 1998; Ergorul & Eichenbaum, 2004; Roberts et al., 2008; Zhou & Crystal, 2009) or by the occasion in which they took place using either temporal or non-temporal identifiers, suggested by WWWhich memory (Eacott & Easton, 2010, 2012; Easton & Eacott, 2008). In spite of their central role within the widely used WWWhen paradigm, humans rarely use specific temporal cues when remembering episodes, relying instead on non-temporal information e.g. information about the weather, people who were there, and environmental context (Friedman, 1993; Wagenaar, 1986; Wells et al. 2014). Clearly, both temporal and non-temporal identifiers of an event can be thought of as sources that specify conditions under which memories were encoded (Johnson et al. 1993), and can be interrogated using standard human memory paradigms.

To date, relatively few attempts have been made to assess the construct validity of the WWWhen or WWWhich paradigms by testing them within humans. Previous research with human participants has shown that while memory for both temporal and contextual information can be accurate (Easton, Webster, & Eacott, 2012; Holland & Smulders, 2011), the two models engage different retrieval processes within the dual process framework (Easton et al., 2012). Dual process theory states that memory can be retrieved using two processes, *recollection* and/or *familiarity* (for reviews, see Vilberg & Rugg, 2008, and Yonelinas, 2002). Within this framework, a memory supported by recollection is retrieved alongside its source, which unambiguously supports the recognition judgement. Familiarity, on the other hand, does not result in source retrieval—merely the awareness that the recognised stimulus relates to something from the past. These processes are often assessed using what is known as a Remember/Know (R/K) procedure in which participants are given a recognition memory test for items and/or sources seen during a study phase (c.f. Jacoby, 1991). For correctly identified items Remember and Know responses are given, where Remember corresponds to the process of recollection and Know to the process of familiarity (Dewhurst et al. 2009). Given this established operationalisation of the dual processes, the only retrieval process capable of supporting source information is recollection. Melding the two discussed sets of frameworks, it should therefore be expected that the temporal source component of WWWhen memories, and the contextual source component of WWWhich memories would both necessarily recruit recollection. This has led some researchers to suggest that recollection is truly episodic, while familiarity, which does not require the retrieval of an integrated memory of multiple features of an event, is not (Clayton & Dickinson, 1998; Easton et al., 2012)[1].

Easton et al. (2012) examined retrieval processes in human participants by projecting images of abstract figures on either a zebra or a checked background on to a large screen in a lecture theatre. They showed that accurate retrieval of contextual information was possible only using recollection. They also showed that, contrary to what would be expected within the dual process framework, memory for the order of stimulus presentation could be retrieved accurately using either recollection or familiarity. This finding questions the WWWhen model's assumptions of the integrity of temporal source cues to episodic memory, at least in humans. Moreover, it suggests that retrieval of temporal sources should not be treated as equivalent to contextual sources within the WWWhich model. There was an alternative explanation for the familiarity-supported temporal source component which is that familiarity judgments can be supported using memory strength (i.e. how little the encoding episode had decayed from memory indicating how recently it was encountered) rather than the retrieval of a true temporal source. We sought to conceptually replicate these findings using a longer sequence of

[1] It is worth noting that this episodic/non-episodic distinction revisits Tulving's (1983) original formulation of recollection as episodic and familiarity as semantic. Human memory researchers have since interpreted both processes as episodic within the dual process framework (e.g. since Yonelinas, 2001), making this an area in which there is little consensus across animal and human memory research.

time points both during study and test (6 time points compared to 2). This allowed us to examine the nature of the errors produced when attempting to use temporal source judgements. If temporal source questions were being solved using memory strength rather than retrieval of the precise temporal source then we would expect incorrect responses to cluster around the position of the correct response within the sequence of presentation at encoding. We additionally used an immersive testing environment to more closely match the conditions under which temporal source has been shown to be integral to episodic memory in animals.

To this end, we assessed time and context retrieval in an immersive virtual environment using an amended version of the standard Remember-Know procedure (Dewhurst et al., 2009; Donaldson et al. 1996), with the only difference being that we asked for 'Familiar' instead of 'Know' responses in order to distinguish between the two familiarity processes suggested by Dewhurst et al. (2009). Over two experiments, we used a paradigm, in which participants moved through, and encountered a series of 3D objects in a virtual environment. We made the environment as immersive as possible by projecting it onto a wall in a darkened room, which was intended to give a strong sense of being present *within* the environment. During the study, participants encountered objects in different weather contexts and at different times (points in a sequence). Their memory for the items and their temporal or contextual sources was assessed alongside a judgement of recollective experience: Remember judgements indicating that participants could retrieve details surrounding the event when the object was presented, indicative of retrieval using recollection; and Familiar judgements indicating that participants only knew that an object had been seen without memory for surrounding information, indicative of retrieval using familiarity. Based on Easton et al. (2012) as well as a recent paper by Saive and colleagues (2015), we predicted that contextual source memory would be more accurate following self-reported engagement of recollection compared to familiarity, but that temporal source memory would be equally accurate across judgements that engaged recollection or familiarity.

2. Materials and Methods

2.1. Experiment 1

2.1.1. Virtual Environment and stimuli

The virtual environment was created using the Valve Hammer World Editor (Valve Software, 2006), was projected onto a 375 cm x 250 cm screen at 1024 x 768 resolution, and was navigated through using a games console controller (Xbox 360, Microsoft). Participants were seated 420 cm from the wall. The virtual environment consisted of two rooms connected by a single doorway: a Start Room, where participants began each trial; and a Main Room where they encountered stimuli, (Figure 1A). The Main Room had three windows facing a Courtyard. The Courtyard contained landmarks consisting of a sculpture, a car and a perimeter of buildings. Context was manipulated by altering the weather conditions in the Courtyard whilst keeping the landmarks constant. Six different contexts were used: sun, snow, lightning, rain, fog, and wind (Figure 1B). Rain used the appearance and sound of rain hitting the windows. Wind used leaves blowing in the courtyard and the sound of howling wind. Lightning used flashes and the sound of thunder. Sun used clear blue skies and bird song. Fog used grey mist and was silent. Snow used falling snowflakes and snow-covered ground and was silent.

A pool of 186 3D stimuli from the Valve Hammer and Garry's mod databases representing a range of everyday objects (see Figure 1A for examples) was used, from which 72 were randomly drawn for each participant. 36 objects were presented in the virtual environment and the recognition test (targets), while the remaining 36 were only presented during the recognition test (lures). Instructions and memory test were displayed on a laptop using Psychophysics Toolbox version 3 (Brainard, 1997; Kleiner et al. 2007) for Matlab (Natick, MA, 2013).

2.1.2. Participants

There were 32 participants (22 female) with an average age of 21.8 years. Four participants stopped the study early, two due to feelings of nausea brought about by the high level of immersion, and two due to technical errors. The results from the remaining 28 (19 female) participants with an average age of 22.0 years were used for analysis. Ethical approval for the study was obtained from the University of St Andrews' University Teaching and Research Ethics Committee. All participants gave informed consent prior to taking part in the study.

2.1.3. Procedure

Each Study Phase consisted of six self-paced trials. During a single trial, participants made their way from the Start Room to the Main Room and located the stimulus behind the barrier. They were instructed to identify the object and judge its size relative to the surrounding furniture and give their response verbally, this response was not recorded but served as a measure to ensure that participants attended to the stimuli. Participants then navigated back to the start room. This procedure was repeated until six objects had been encountered. Each time participants entered the Main Room a new object was presented and the context was changed (randomised according to a Latin Square design). Within each Study Phase, each stimulus was therefore associated with one context and one position in the sequence (1-6).

Each Test Phase consisted of 12 self-paced trials. Memory for the information seen in the virtual environment was assessed using four different questions per trial. First, to assess object recognition participants were first shown an image of an item on the screen and asked whether it was 'old' (seen in the Virtual Environment; indicated using the '1' key) or whether it was 'new' (not seen in the Virtual Environment; indicated using the '0' key).

Following the object recognition participants were asked about the source of the item. Source questions either asked for the context the item was seen in or when in the presentation sequence the item was seen. There was a 50:50 ratio of context and sequence question in each test phase and the questions were randomised within participants. For both questions all response options were displayed on the screen simultaneously. Responses were given using the keys 1-6 on the keyboard. The order of the context response alternatives during the test phase did not map on to the sequence in which they were presented during study. This was to reduce the probability of participants using temporal information to retrieve context, and vice versa. Source questions were asked regardless of whether participants responded that the item was old or new. This was to allow for the analysis of source judgments in the absence of object recognition. If participants were certain that an item was new they were instructed to pick a source response at random. Participants subsequently rated their confidence in their source response on a scale from 1 (low confidence) to 3 (high confidence). Finally, having participants give a Remember (R), Familiar (F) or Guess (G) response, where Remember corresponded to recollection and Familiar to familiarity, assessed recollective experience. This question was asked last so as not to influence confidence assignment (Williams et al. 2013). Definitions for these judgements, adapted from Gardiner et al. (1996), were provided at the beginning of the experiment. Participants also had the option to review these definitions before each new test phase:

- *Remember - Recognition of the object brings back details of the experience when the object was encountered in the virtual world.*
- *Familiar - Knowing that the object was encountered because of a feeling of familiarity, without recollecting the particular occurrence.*
- *Guess - A decision strategy used when object information does not elicit either the experience of remembering or that of familiarity.*

There were six Study-Test Phases giving a total of 72 trials: 36 target trials (18 context trials and 18 sequence trials); and 36 lure trials.

2.2. Experiment 2

2.2.1. Virtual Environment and stimuli

The virtual environment was altered in Experiment 2 to create three stalls along the short wall of the main room in which the objects were placed (see Figure 1C and 1D).

Items were drawn from the same pool as in Experiment 1, with 144 items randomly selected for each participant. 108 of these were seen in the Virtual Environment and 36 acted as lures in the test phase.

2.2.2. Participants

29 participants (20 female) with an average age of 22.7 were recruited. 5 participants stopped the study early due to technical errors. This left 24 participants (16 female) with an average age of 23.0 years whose data were analysed.

2.2.3. Procedure

The Study Phase in Experiment 2 followed the same structure as in Experiment 1, with the exception that three objects were seen in each trial, which across 6 study-test blocks of 6 trials gave 108 to-be-remembered objects. In each trial an object was placed in each of the 3 individual stalls with a barrier in front of each one. Participants were instructed to navigate to one stall at a time, on arrival at the stall the barrier would disappear and the object would be visible for five seconds before disappearing (Figure 3).

Each Test Phase consisted of 12 self-paced trials, only 1 of the 3 to-be-remembered items from each study trial was presented. Aside from this the Test Phase was identical to that in Experiment 1. There were 6 study-test blocks resulting in a total of 36 target trials (18 context and 18 sequence trials) and 36 new lure trials over the entire experiment.

<u>2.3. Statistics</u>

Due to some missing values, a linear mixed model (LMM) was used to analyse the effect of source (context/sequence) and recollective experience (Remember/Familiar) on accuracy of retrieval in both Experiment 1 and 2. The LMM was selected by fitting models with different covariance structures and choosing the most appropriate based on the model with the lowest Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC). The covariance structures assessed were first-order autoregressive (AR1), unstructured, and compound symmetry. The LMM had source (context/sequence) and recollective experience (Remember/Familiar judgments) as repeated measures. Source and recollective experience were also entered as fixed factors along with a source by recollective experience interaction.

The model evaluation found that the AR1 covariance structure resulted in the lowest AIC and BIC values compared to compound symmetry and unstructured covariance structures, where lower AIC and BIC values indicate a better model fit (see Table 1 and 2 in Supplementary Materials for a comparison of covariance structures).

Accuracy was further compared to chance level using one-sample *t*-tests in order to see whether performance in the context or sequence would be significantly above what would be expected if participants were merely guessing. Chance level performance was calculated as 1/6, as in each trial there was a 1 in 6 chance of getting the correct answer by simply guessing.

To examine the possibility that retrieval of temporal information using familiarity was due to trace strength we examined serial position curves, plotting the proportion of error responses for both sequence and context questions. The aim was to characterise each source retrieval task as operating via a discrete threshold process, i.e. with randomly distributed errors, or via continuous trace strength process, i.e. error rates gradually dropping off as with increasing dissimilarity to the correct response option. For the sequence condition, this was carried out by examining the proportion of responses around items presented in the 3rd and 4th sequence positions (3rd and 4th sequence position items, those in the middle of lists, were chosen to provide maximal opportunity to examine the tails of the serial position curves). For the context condition, serial position of incorrect responses was ordered according to frequency of incorrect response within each context. These distributions were then averaged across all contexts. It should be noted that the serial position ordering of the context condition items was carried out to give the greatest opportunity to uncover a gradual drop-off serial position curve, should one be present in the data.

The proportion of error responses were compared to chance levels using one-sample t-tests in order to see whether the frequency with which they were made differed from what would have been expected if participants were making errors at random. Here chance levels were calculated as 1 minus the proportion of correct responses, divided by the number of incorrect response options to give a level at which all errors could be distributed evenly.

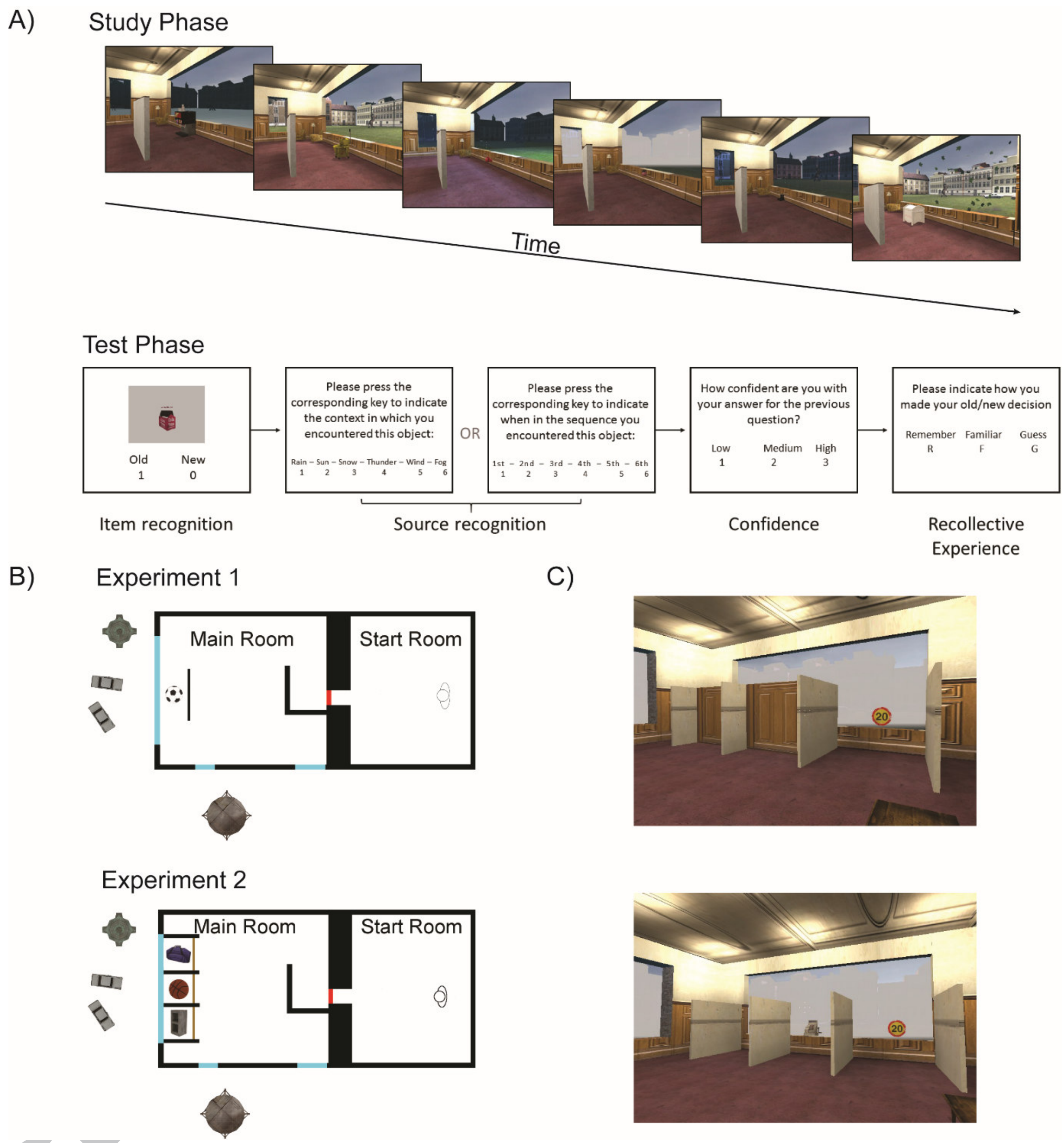


Figure 1. A) Schematic of the Study and Test Phase in both experiments. Images in the Study Phase are from Experiment 1 where only one object was presented per trial. **B)** Overviews of the virtual environment and the outside landmarks for both Experiment 1 (top) and Experiment 2 (bottom). **C)** Screenshots showing the stalls in which objects were found in Experiment 2. The top image shows one stall being opened while the other two remain unexplored. Bottom image shows all stalls being open with the images behind the panels being shown.

3. Results

3.1. Experiment 1

3.1.1. Object Recognition

Object recognition accuracy was high, with the proportion of correctly identified targets (hits) and correctly identified lures (correct rejections) at .94 (SD=0.13) and 0.95 (SD=0.14) respectively. The proportion of R, F, and G responses was similar across context and sequence questions. R responses followed context hits at a rate of 0.81 (SD=0.29), with F at 0.15 (SD=0.26) and G at 0.04 (SD=0.07). For sequence questions, R responses succeeded sequence hits at a rate of 0.86 (SD=0.26) F at 0.11 (SD=0.20) and G at 0.03 (SD=0.09). G responses, included to remove contamination of R and F responses by actual guesses (Gardiner et al., 1996), were not analysed any further. A 2 (recollective experience: R/F) x2 (source: context/sequence) ANOVA comparing the R and F response rates for context and sequence, presented above, found a main effect of recollective experience, $F_{(1,27)}$=65.20, p<0.001, $\eta_p{}^2$=0.71, meaning that R responses were given at a significantly higher rate than F responses. The main effect of source was nonsignificant, $F_{(1,27)}$=0.40, p=0.84, $\eta_p{}^2$=0.001, as was the recollective experience by source interaction, $F_{(1,27)}$=1.46, p=0.24, $\eta_p{}^2$=0.05, suggesting that the R and F response rates were not significantly different across context and sequence questions.

3.1.2. Source memory by recollective experience

For questions about contextual information, participants were more accurate when giving R judgments, 0.68 (SD=0.25), compared to when giving F judgments, 0.21 (SD=0.19). Responses for sequence questions followed a similar pattern with higher accuracy for R judgments, 0.74 (SD=0.23), than for F judgments, 0.43 (SD=0.33). The LMM showed that there was a significant effect of recollective experience (R/F judgments), β=-0.37, *SE*=0.79, p<0.001, indicating that source questions followed by R judgments were more accurate compared to F judgments across both context and sequence. The effect of source (context/sequence) was not significant, β=-0.57, *SE*=0.55, p=0.30, meaning that accuracy was not different between context and sequence questions. Nor was the source by recollective experience interaction significant, β=-0.11, *SE*=0.10, p=0.26.

We hypothesised that accurate retrieval of contextual information could only be achieved using recollection and not familiarity, whereas sequence information could be retrieved accurately using both recollection and familiarity. However, to fully assess this, performance using the different retrieval strategies must be compared to chance performance to ensure that participants could accurately retrieve source information (Figure 2). Comparing performance against chance levels using one-sample *t*-tests it was found that for context questions, performance was only above chance for R, $t_{(25)}$=10.37, p<0.001, d=2.01, but not for F, $t_{(15)}$=0.90, p=0.38, d=0.08. On the other hand, performance was above chance for both R, $t_{(27)}$=12.99, p<0.001, d=1.85, and F for sequence questions, $t_{(11)}$=2.82, p=0.02, d=0.80. This analysis indicates that context questions could only be answered above chance levels when based on recollection-based retrieval, whereas sequence judgments can be made correctly using either recollection or familiarity.

3.1.3. Trace strength

In the trace strength analysis for sequence errors (Figure 3A) in Experiment 1 only the proportion of errors for pos-1 (that is one sequential position behind the correct response) were significantly higher than chance level, $t_{(27)}$=5.21, p<0.001. This indicates some forward asymmetry in error judgments where participants tended to make errors that were one step ahead of the correct option in the sequence. Analysis of context errors (Figure 3B) showed that the proportion of errors for Rank 1, which is the average of the most common error made for each context question, was above chance level, $t_{(27)}$=2.26, p=0.032. None of the other ranks were above chance.

3.1.4. Confidence judgments

Examining confidence judgments, R judgments tended to be given high levels of confidence at 2.72 (SD=0.29) for context and 2.76 (SD=0.30) for sequence questions. F judgments, on the other hand, had somewhat lower confidence levels with 1.96 (SD=0.30) for context and 2.00 (SD=0.44) for sequence questions. A 2 (recollective experience: R/F) x2 (source: context/sequence) ANOVA revealed a significant effect of recollective experience , $F_{(1,8)}$=49.59, p<0.001, η_p^2=0.86, while no significant effect was found for source , $F_{(1,8)}$=0.48, p=0.51, η_p^2=0.05, nor the source by recollective experience interaction, $F_{(1,8)}$=0.002, p=0.96, η_p^2<0.001. This suggests that participants were making R judgments with significantly higher confidence compared to F judgments, independent of source.

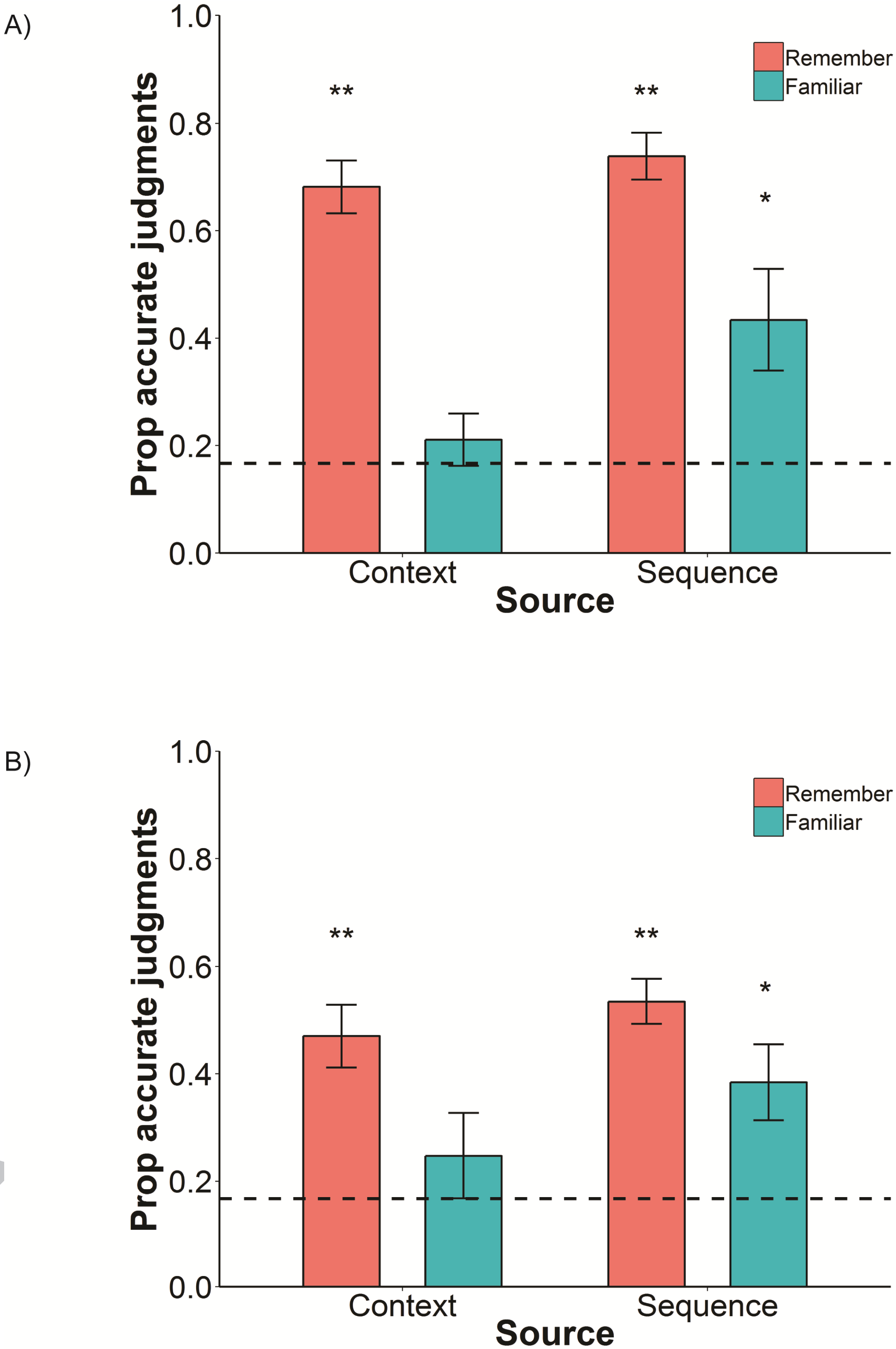


Figure 2. Context and sequence accuracy by R and F judgments in A) Experiment 1, and B) Experiment 2. Error bars represent standard error of the mean. ** indicates $p \leq 0.001$ and * indicates $p \leq 0.05$ compared to chance level performance (dotted line, 0.1667).

3.2. Experiment 2

3.2.1. Object recognition

Experiment 2 addressed the high object recognition scores and relatively low proportion of Familiar responses in Experiment 1. Object discrimination remained relatively high, with hits at 0.85 (SD=0.19), and correct rejections at 0.92 (SD=0.18). Similar to Experiment 1, proportions of R-F-G responses for correct source judgments were roughly equal following context hits, with R judgments being given 0.89 (SD =0.17) of the time and F 0.11 (SD =0.17), and sequence hits, where R judgments had a rate of 0.84 (SD =0.16) and F at 0.16 (SD =0.16). There were no G judgments for either context or sequence questions in Experiment 2. As there were no G judgments, the proportions of R and F judgments were not independent meaning that an ANOVA could not be carried out. Instead a paired sample t-test was used to compare the proportion of R responses between context and sequence questions. This t-test was non-significant, $t_{(22)}$=1.49, p=0.15, indicating that the proportions of R judgments given to context and sequence questions were not different, and in extension neither were the proportions of F judgments for context and sequence significantly different.

3.2.2. Source memory by recollective experience

Context accuracy was higher when justified by R judgments, 0.47 (SD=0.29) compared to when justified by F judgments, 0.25 (SD=0.34). The same pattern, although attenuated, was seen for sequence questions where accuracy followed by R judgments were 0.53 (SD=0.21) relative to an accuracy of 0.38 (SD=0.32) for sequence questions followed by F judgments. Results from the LMM revealed that there was no significant effect of recollective experience in Experiment 2, β=-0.15, SE=0.82, p=0.07, suggesting that accuracy was not difference between R and F judgments across context and sequence questions. Furthermore, the effect of source, β=-0.06, SE=0.69, p=0.35, as well as the source by recollective experience interaction, β=-0.07, SE=0.11, p=0.48, were non-significant.

As in Experiment 1, accurate retrieval of source information was examined using one sample t-tests. These demonstrated that for context questions performance was above chance for R, $t_{(23)}$=5.16, p<0.001, d=1.05, but not for F, $t_{(17)}$=1.01, p=0.33, d=0.24. For time questions, performance was above chance for both R, $t_{(23)}$=8.70, p<0.001, d=1.77, and F, $t_{(19)}$=3.07, p=0.006, d=0.67. This again demonstrates that participants could only answer context questions accurately using recollection, but not familiarity. Sequence questions, on the other hand, could be answered accurately using both recollection and familiarity.

3.2.3. Trace strength

In the trace strength analysis In Experiment 2, the proportion of errors were above chance for pos1, $t_{(23)}$=2.62, p=0.015, and neg1, $t_{(23)}$=3.35, p=0.003. Here errors cluster around the correct temporal position which could argue for the use of trace strength to retrieve temporal information. Having a general idea of when in the sequence an item was presented this could narrow the correct temporal accuracy down to a scale of +1 to -1 position, in line with trace strength spanning the most temporal proximal positions.

For context errors in Experiment 2, only the proportion of error judgments for Rank 1 was above chance level, $t_{(23)}$=2.34, p=0.028. Similar to the sequence errors, the proportions of context errors were not randomly distributed across all response options. Instead incorrect responses tended to cluster in the most common error response category.

3.2.4. Confidence levels

Confidence judgments, R judgments for context and sequence were 2.43 (SD=0.58) and 2.39 (SD=0.46) respectively. For F judgments confidence levels were lower at 1.55 (SD=0.36) for context questions and 1.64 (SD=0.41) for sequence. A 2 (recollective experience: R/F) x2 (source: context/sequence) ANOVA was run, again showing a main effect of recollective experience, $F_{(1,14)}$=61.08, p<0.001, η_p^2=0.81, but not for source, $F_{(1,14)}$=0.27, p=0.61, η_p^2=0.02, nor the interaction between source and recollective experience, $F_{(1,14)}$=0.87, p=0.36, η_p^2=0.06. As in Experiment 1 this demonstrates that R judgments were overall more confident than F judgments across both context and sequence questions.

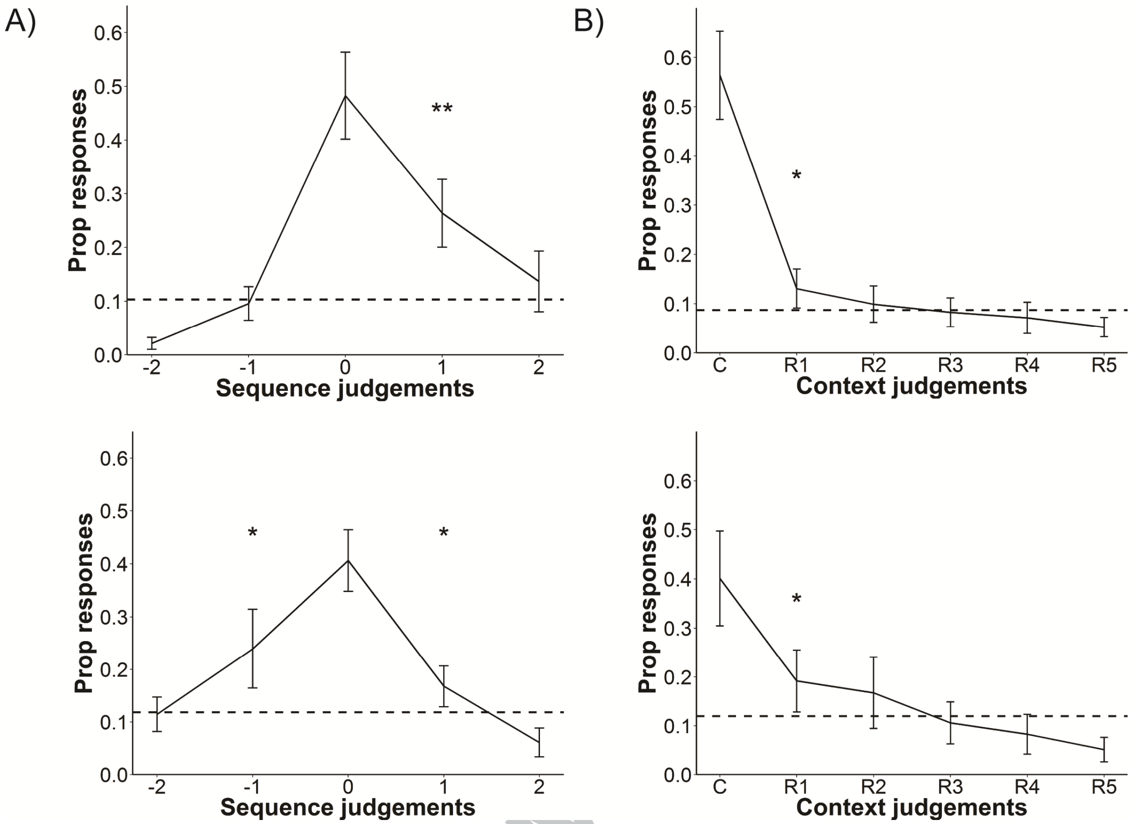


Figure 3. A) Proportion of error responses for sequence judgments for Experiment 1 (top) and Experiment 2 (bottom). 0 corresponds to correct responses, while positive and negative integers indicate the temporal position of responses after or before the correct response. B) Proportion of error responses for context questions in Experiment 1 (top) and Experiment 2 (bottom). C indicates the proportion of correct responses across contexts. Rank 1 (R1) corresponds to the average of the most frequent error for each type of context, Rank 2 (R2) is the average of the 2nd most frequent error, etc. Dotted lines indicate chance level. Error bars represent confidence intervals.

4. Discussion

Our hypothesis, based on the operationalisation of recollection within the dual process model (Yonelinas, 2001) and a previous study examining retrieval processes underpinning the WWWhich and WWWhen paradigms (Easton et al., 2012; Saive et al., 2015), was that accurate memory for objects in context would only be possible using a recollection strategy, while accurate memory for objects within a sequence would be supported by either recollection or familiarity. Results from the two experiments presented here support this hypothesis. We demonstrated that while context can only be accurately retrieved using recollection (the retrieval process underpinning R responses), temporal sequence information can be retrieved accurately using either recollection or familiarity (the retrieval process underpinning F responses).

There are two ways in which these data can be interpreted. The first is that the time an event took place (here defined as its place in a sequence) cannot be used as a source as this would mean it should only be possible to retrieve that information using recollection. This would have implications for using temporal information as a basis for the definition of episodic memory and for using the WWWhen paradigm to model it in animals. Alternatively these data could show that familiarity can be used to retrieve some kinds of source information (time/sequence) and as such should be considered to be as 'episodic' as recollection.

While these results do not discount time as a valid source in episodic memory as it can be retrieved using recollection, they do highlight that there are instances in which the retrieval of temporal information violates the assumptions of source memory under the dual process theory, where memory for a source is argued to rely on recollection alone (Yonelinas, 1999). What is clear from our data is that context information does conform to the accepted view of source information in that it can only be retrieved using recollection and as such this supports a more inclusive model of episodic memory in which events are defined by their occasion, including both time and context (Eacott & Easton, 2010; Eacott & Norman, 2004), rather than by temporal information alone as suggested by what-where-when memory (Cheke & Clayton, 2010; Eichenbaum & Fortin, 2003).

These results can also be viewed in relation to the unitization of stimulus-source associations. In some instances an item and its source, can be encoded as a unit and be retrieved without reference to other units using familiarity. However, for other events sources are encoded as separate units from the item and so memory of the event requires an association to be made between item and source. Retrieval of this integrated memory requires recollection (Diana et al. 2011; Diana et al. 2008; Yonelinas, 1999). In the case of the present study, item-context associations might have been encoded through the latter process, where weather conditions were encoded as a separate unit to the object. On the contrary, sequence information might instead be encoded as a feature of the item itself similar to colour and size, thus enabling the use of familiarity in correct retrieval (Yonelinas, 1999, 2002). It has been suggested that some associative memory might be influenced by a degree of familiarity, such as temporal order judgments relying on the trace strength of an item to be placed in time (Yonelinas, 2002). While we asked participants for specific temporal judgments (i.e. the exact temporal position an item was presented) rather than a relative judgment, it is unlikely that our task could have been solved solely by the use of trace strength. Trace strength could, however, have contributed towards an accurate judgment by distinguishing the correct option from more distant sequential positions and in that way narrowed down the window of plausible responses. Some evidence supporting this was found in the trace strength analysis where it was shown that errors in temporal judgments tended to cluster around the correct response, and that the proportions of such error judgments were significantly greater than expected by chance. The use of familiarity to accurately retrieve sequential information might pose a problem for theories of temporal memory where time is seen as an entirely separate component to item and space (Eichenbaum & Fortin, 2003; Kraus et al., 2013). The current findings suggest that in some circumstances time might not be encoded separately from the object or event, but rather as another feature.

Another interesting finding is that participants did not use temporal information supported by familiarity to retrieve contexts, for example by knowing that an item was the third in the sequence and that snow was the third context. This is consistent with Easton et al. who showed that even when participants only need to remember the order of two contexts they do not combine this with familiarity supported temporal information to solve the task. This is perhaps not surprising as this strategy would be much harder to use in the current experiments as participants would need to remember the order of the 6 context presentations at encoding. However, this reinforces the finding that familiarity supported temporal judgement are not combined with other information to help remember specific episodes.

It has recently been suggested that recollection of associations between objects and the contexts in which they were experienced is aided by semantic links between the two (Saive et al., 2015). For example, recollection of finding a beach ball on a beach may be aided by the semantic link. However, semantic information cannot be used to recall the time at which (or place within a sequence) the beach ball was experienced. In the current study this could mean that semantic links between the objects and the contexts would make retrieval of contextual sources easier. This was not found to be the case, however, as participants' ability to accurately recollect time at which an object was experienced was better than their ability to recollect the context in which objects were experienced suggesting that semantic links to context do not increase participants' ability to recognise an object.

A further consideration is what our findings tell us about the use of time to disambiguate previously experienced events. Counter to the widespread acceptance in the human cognitive literature of familiarity and recollection as complementary episodic retrieval processes (Diana, Yonelinas, & Ranganath, 2007; Yonelinas, 1994, 2002), some authors have suggested that familiarity is a less episodic process than recollection (Easton et al., 2012; Holland & Smulders, 2011). According to this view, the fact that tasks using time to disambiguate events can be solved using familiarity would call into question the validity of using time in tests of episodic memory. However, one important consideration here is that there are many ways to think of and define time in relation to memory (Crystal, 2010; Friedman, 1993). In the present study we have used place in a sequence to define time of an event but other studies have utilised measures of time, such as how long ago (Clayton & Dickinson, 1998; Roberts et al., 2008) or what time of day an event took place (Zhou & Crystal, 2009). While it is clear that place in a sequence can be remembered using familiarity it could be that these other measures of time might require recollection for accurate retrieval.

Nevertheless, sequence memory has successfully been implemented in models of episodic-like memory (Ergorul & Eichenbaum, 2004), and been demonstrated to rely on the hippocampus (Agster, Fortin, & Eichenbaum, 2002; Devito & Eichenbaum, 2011; Fortin, Agster, & Eichenbaum, 2002), which has been shown to be a critical part of the network that processes episodic memory (Burgess, Maguire, & O'Keefe, 2002; Squire & Zola-Morgan, 1991; Vargha-Khadem, 1997). Data from the current experiment might help to explain these findings. Our data show that while sequence memory can be solved using familiarity the most accurate and common process used to recall sequence information is recollection. Given that the hippocampus has been implicated in accurate use of recollection (Fortin, Wright, & Eichenbaum, 2004; Sauvage, Fortin, Owens, Yonelinas, & Eichenbaum, 2008; Diana et al., 2007; for review see Eichenbaum et al., 2007), it is not surprising that humans and animals with damage to the hippocampus are impaired at remembering position within a sequence. Damage to the hippocampus would remove participants' ability to use recollection to solve the task resulting in a deficit. However, these participants would still be able to use familiarity to solve the task even though it is less accurate. This is consistent with the fact that animals with hippocampal lesions are above chance in some measures of sequence memory even though they are impaired relative to controls (Fortin et al. 2002).

5. Conclusions

The current data support the hypothesis that the contextual features of an event can only be retrieved from memory using a recollection strategy. This is consistent with context information being used as a source for episodic memory and supports a more inclusive definition that does not rely solely on temporal information to disambiguate events. They also suggest that while temporal information can be used as a source in some circumstances there are occasions in which it can be retrieved using a strategy based on familiarity. This could be interpreted as showing that time is not always a source, that time can be encoded as an intrinsic feature of an item or that familiarity can sometimes be used to retrieve source information.

Acknowledgements

The authors report no conflicts of interest. Participant payment was provided by the School of Psychology and Neuroscience ResPay scheme.

References

Agster, K. L., Fortin, N. J., & Eichenbaum, H. (2002). The hippocampus and disambiguation of overlapping sequences. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *22*(13), 5760–8. http://doi.org/20026559

Babb, S. J., & Crystal, J. D. (2005). Discrimination of what, when, and where: Implications for episodic-like memory in rats. *Learning and Motivation*, *36*(2), 177–189. http://doi.org/10.1016/j.lmot.2005.02.009

Babb, S. J., & Crystal, J. D. (2006). Episodic-like memory in the rat. *Current Biology : CB*, *16*(13), 1317–21. http://doi.org/10.1016/j.cub.2006.05.025

Burgess, N., Maguire, E. a, & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*(4), 625–41. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12194864

Cheke, L. G., & Clayton, N. S. (2010). Mental time travel in animals. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(6), 915–930. http://doi.org/10.1002/wcs.59

Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, *395*(September), 272 – 274.

Crystal, J. D. (2010). Episodic-like memory in animals. *Behavioural Brain Research*, *215*(2), 235–43. http://doi.org/10.1016/j.bbr.2010.03.005

Davis, K. E., Easton, A., Eacott, M. J., & Gigg, J. (2013). Episodic-like memory for what-where-which occasion is selectively impaired in the 3xTgAD mouse model of Alzheimer's disease. *Journal of Alzheimer's Disease : JAD*, *33*(3), 681–98. http://doi.org/10.3233/JAD-2012-121543

Devito, L. M., & Eichenbaum, H. (2011). Memory for the order of events in specific sequences: contributions of the hippocampus and medial prefrontal cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(9), 3169–75. http://doi.org/10.1523/JNEUROSCI.4202-10.2011

Dewhurst, S. A., Conway, M. A., & Brandt, K. R. (2009). Tracking the R-to-K shift: Changes in memory awareness across repeated tests. *Applied Cognitive Psychology*, *23*(6), 849–858. http://doi.org/10.1002/acp.1517

Diana, R. A., Van den Boom, W., Yonelinas, A. P., & Ranganath, C. (2011). ERP correlates of source memory: unitized source information increases familiarity-based retrieval. *Brain Research*, *1367*, 278–86. http://doi.org/10.1016/j.brainres.2010.10.030

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends in Cognitive Sciences*, *11*(9), 379–86. http://doi.org/10.1016/j.tics.2007.08.001

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). The effects of unitization on familiarity-based source memory: testing a behavioral prediction derived from neuroimaging data. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*(4), 730–40. http://doi.org/10.1037/0278-7393.34.4.730

Donaldson, W., Mackenzie, T. M., & Underhill, C. F. (1996). A comparison of recollective memory and source monitoring. *Psychonomic Bulletin & Review*, *3*(4), 486–90. http://doi.org/10.3758/BF03214551

Eacott, M. J., & Easton, A. (2010). Episodic memory in animals: remembering which occasion. *Neuropsychologia*, *48*(8), 2273–80. http://doi.org/10.1016/j.neuropsychologia.2009.11.002

Eacott, M. J., & Easton, A. (2012). Remembering the past and thinking about the future: Is it really about time? *Learning and Motivation*, *43*(4), 200–208.

http://doi.org/10.1016/j.lmot.2012.05.012

Eacott, M. J., & Norman, G. (2004). Integrated memory for object, place, and context in rats: a possible model of episodic-like memory? *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *24*(8), 1948–53. http://doi.org/10.1523/JNEUROSCI.2975-03.2004

Easton, A., & Eacott, M. J. (2008). A new working definition of episodic memory : replacing "'when'" with "'which.'" *Handbook of Episodic Memory*, *18*(08). http://doi.org/10.1016/S1569-7339(08)00211-7

Easton, A., Webster, L. A. D., & Eacott, M. J. (2012). The episodic nature of episodic-like memories. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *19*(4), 146–50. http://doi.org/10.1101/lm.025676.112

Eichenbaum, H., & Fortin, N. (2003). Episodic memory and the hippocampus: it's about time. *Current Directions in Psychological Science*, *12*(2), 53–57. http://doi.org/10.1111/1467-8721.01225

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience*, *30*, 123–52. http://doi.org/10.1146/annurev.neuro.30.051606.094328

Ergorul, C., & Eichenbaum, H. (2004). The Hippocampus and Memory for " What ," " Where ," and " When ." *Learning & Memory (Cold Spring Harbor, N.Y.)*, *11*(13), 397–405. http://doi.org/10.1101/lm.73304.mals

Fortin, N. J., Agster, K. L., & Eichenbaum, H. B. (2002). Critical role of the hippocampus in memory for sequences of events. *Nature Neuroscience*, *5*(5), 458–62. http://doi.org/10.1038/nn834

Fortin, N. J., Wright, S. P., & Eichenbaum, H. (2004). Recollection-like memory retrieval in rats is dependent on the hippocampus. *Nature*, *431*(7005), 188–91. http://doi.org/10.1038/nature02853

Friedman, W. J. (1993). Memory for the time of past events. *Psychological Bulletin*, *113*(1), 44–66.

Gardiner, J. M., Java, R. I., & Richardson-Klavehn, A. (1996). How level of processing really influences awareness in recognition memory. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *50*(1), 114–122. http://doi.org/10.1037/1196-1961.50.1.114

Holland, S. M., & Smulders, T. V. (2011). Do humans use episodic memory to solve a What-Where-When memory task? *Animal Cognition*, *14*(1), 95–102. http://doi.org/10.1007/s10071-010-0346-5

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513–541. http://doi.org/10.1016/0749-596X(91)90025-F

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source Monitoring.pdf. *Psychological Bulletin*, *114*, 3–28.

Jozet-Alves, C., Bertin, M., & Clayton, N. S. (2013). Evidence of episodic-like memory in cuttlefish. *Current Biology : CB*, *23*(23), R1033–5. http://doi.org/10.1016/j.cub.2013.10.021

Kart-Teke, E., De Souza Silva, M. a, Huston, J. P., & Dere, E. (2006). Wistar rats show episodic-like memory for unique experiences. *Neurobiology of Learning and Memory*, *85*(2), 173–82. http://doi.org/10.1016/j.nlm.2005.10.002

Kraus, B. J., Robinson, R. J., White, J. A., Eichenbaum, H., & Hasselmo, M. E. (2013). Hippocampal "time cells": time versus path integration. *Neuron*, *78*(6), 1090–101.

http://doi.org/10.1016/j.neuron.2013.04.015

Langston, R. F., & Wood, E. R. (2010). Associative recognition and the hippocampus: differential effects of hippocampal lesions on object-place, object-context and object-place-context memory. *Hippocampus*, *20*(10), 1139–53. http://doi.org/10.1002/hipo.20714

Martin-Ordas, G., Berntsen, D., & Call, J. (2013). Memory for distant past events in chimpanzees and orangutans. *Current Biology : CB*, *23*(15), 1438–41. http://doi.org/10.1016/j.cub.2013.06.017

Roberts, W. A., Feeney, M. C., Macpherson, K., Petter, M., McMillan, N., & Musolino, E. (2008). Episodic-like memory in rats: is it based on when or how long ago? *Science (New York, N.Y.)*, *320*(5872), 113–5. http://doi.org/10.1126/science.1152709

Saive, A.-L., Royet, J.-P., Garcia, S., Thévenet, M., & Plailly, J. (2015). "What-Where-Which" Episodic Retrieval Requires Conscious Recollection and Is Promoted by Semantic Knowledge. *PloS One*, *10*(12), e0143767. http://doi.org/10.1371/journal.pone.0143767

Sauvage, M. M., Fortin, N. J., Owens, C. B., Yonelinas, A. P., & Eichenbaum, H. (2008). Recognition memory: opposite effects of hippocampal damage on recollection and familiarity. *Nature Neuroscience*, *11*(1), 16–8. http://doi.org/10.1038/nn2016

Squire, L., & Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science*, *253*(5026), 1380–1386. http://doi.org/10.1126/science.1896849

Suddendorf, T., & Busby, J. (2003). Mental time travel in animals? *Trends in Cognitive Sciences*, *7*(9), 391–396. http://doi.org/10.1016/S1364-6613(03)00187-6

Tulving, E. (2002). Episodic Memory: From Mind to Brain. *Annual Review of Psychology*, *53*, 1–25.

Vargha-Khadem, F. (1997). Differential Effects of Early Hippocampal Pathology on Episodic and Semantic Memory. *Science*, *277*(5324), 376–380. http://doi.org/10.1126/science.277.5324.376

Vilberg, K. L., & Rugg, M. D. (2008). Memory retrieval and the parietal cortex: a review of evidence from a dual-process perspective. *Neuropsychologia*, *46*(7), 1787–99. http://doi.org/10.1016/j.neuropsychologia.2008.01.004

Wagenaar, W. A. (1986). My Memory : A Study of Autobiographical over Six Years. *Cognitive Psychology*, *18*, 225–252.

Wells, C., Morrison, C. M., & Conway, M. a. (2014). Adult recollections of childhood memories: What details can be recalled? *Quarterly Journal of Experimental Psychology (2006)*, *67*(7), 1249–61. http://doi.org/10.1080/17470218.2013.856451

Williams, H. L., Conway, M. a, & Moulin, C. J. a. (2013). Remembering and knowing: using another's subjective report to make inferences about memory strength and subjective experience. *Consciousness and Cognition*, *22*(2), 572–88. http://doi.org/10.1016/j.concog.2013.03.009

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *20*(6), 1341–54. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7983467

Yonelinas, A. P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: a formal dual-process model and an analysis of receiver operating characteristics. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *25*(6), 1415–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10605829

Yonelinas, A. P. (2001). Components of episodic memory: the contribution of recollection and familiarity. *Philosophical Transactions of the Royal Society of London. Series B, Biological*

Sciences, *356*(1413), 1363–74. http://doi.org/10.1098/rstb.2001.0939

Yonelinas, A. P. (2002). The Nature of Recollection and Familiarity: A Review of 30 Years of Research. *Journal of Memory and Language*, *46*(3), 441–517. http://doi.org/10.1006/jmla.2002.2864

Zhou, W., & Crystal, J. D. (2009). Evidence for remembering when events occurred in a rodent model of episodic memory. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(23), 9525–9. http://doi.org/10.1073/pnas.0904360106

Disambiguating past events: Accurate source memory for time and context depends on different retrieval processes, NLM-15-378

Highlights:

Context of encoding events can only be retrieved from memory using recollection.

Time of encoding events can be retrieved using either recollection or familiarity.

Familiarity based time judgements are based on memory trace strength.