

# Part or Parcel? Contextual Binding of Events in Episodic Memory

*Iris Trinkler, John King, Hugo Spiers#, Neil Burgess\**

Institute of Cognitive Neuroscience, and  
Department of Anatomy and Developmental Biology  
University College London,  
17 Queen Square, London WC1N 3AR, U.K.

\* Corresponding author. E-mail: n.burgess@ucl.ac.uk.

# Present address: Functional Imaging Lab., Wellcome Department of Imaging Neuroscience, University College London, 12 Queen Square, London WC1N 3BG

Acknowledgements: This research was supported by an MRC senior research fellowship to NB, UCL Graduate school and ORS studentships to IT and an MRC studentship to HS. We thank Rebecca Dukeson for help collecting data and Tom Hartley and John Morton for many useful discussions and for commenting on an early version of this manuscript. We also thank the editors and an anonymous referee for helpful comments.

## ***Introduction***

In our daily life we experience a vast numbers of events involving objects, people, and places. Memory for personally experienced events is often referred to as episodic memory and has been distinguished from semantic memory, memory for factual information, by the fact that episodic memories contain both information about what happened and a specific spatial and temporal context (Tulving, 1972). The "prototypical unit of an episodic memory is an event" has been suggested by Tulving (1983, p.223), and the different elements of an event are thought to be strongly tied together to provide a single encapsulated unit, allowing "re-experience" of all aspects of the event at retrieval (Tulving 2002, 1983, 1972).

In this chapter we examine whether events are the units of episodic memory. Taken at face value this would imply that episodic memory is holistic: when an event is remembered, *all* of its elements including the spatio-temporal context are remembered together. On the other hand if this viewpoint were completely untrue, memory for different elements of an event may be remembered or forgotten *independently*. In between these theoretical extremes, we might characterize the argument that episodic memory for events is holistic in terms of the size of the correlation between performance when an event is retrieved via one cue and performance when the same event is retrieved via another cue. A fully holistic view would predict maximal correlation. A fully fragmented view or 'independent model' of memory for the many types of association comprising an event would predict no such correlation.

Support for the different interpretations (holistic or fragmentary) can be found in previous research. The holistic view is strongly implied by Tulving's theoretical stand point and that of his co-authors. For example: Episodic remembering is "the kind of awareness that characterizes 'mental re-living' of happenings from one's personal past. It is phemenologically known to all healthy people who can 'travel back in time in their own minds' " Duzel et al., (1997, p5973). Tulving and his colleagues are not alone in their suggestion that events might be the units of episodic memory. Fisher and Chandler (1991) remark that the episodic memory system "treats information in a close temporal-

spatial proximity as an event that is represented in an *isolated* trace. Later activation of that trace produces recollection of that specific event” (p.722, our italics), based on observed interdependence between the recall of different event sets. A study by Brewer and Dupree (1983) suggests that, for at least some types of events, recall appears to be all or none. In their experiments participants were shown movies in which actors performed goal directed actions. In some cases there was a causal link between elements in the event and in others the link was solely temporal. Recall of the causally related events tended to be all or none, while the recall of the non-causally related events tended to be less well correlated.

Jones (1976) examined the recall of different elements of an event using one or several elements as retrieval cues. Participants observed sequences of pictures of colored objects, each in a specific location within the scene. They were then cued with the color, shape, spatial position or sequential position, or with combinations of these cues, and their ability to retrieve the remaining elements was tested. Jones noticed several patterns in this data. Firstly, the nature of sequential position as a memory cue was different from the other elements in being asymmetrical. Retrieval of sequential position was poor compared to its usefulness as a retrieval cue (in addition, retrieval of serial position decreased with serial position while retrieval of other information from serial position showed a u-shaped curve, being best near to the start or end of a sequence). By contrast other elements were used symmetrically: retrieval of element A by element B being as good as retrieval of element B by element A. This symmetrical use of cues is also proposed by Asch and Ebenholz (1962). Secondly, recall performance was not found to increase dramatically with additional cues, ruling out a fully independent model in which each pair-wise association contributes independently to the probability of success. Jones suggested that memories of the visual characteristics of his events (object, color and location, i.e. ignoring sequential position) were stored as independent but holistic fragments. Thus, those elements represented within the same fragment would act holistically (all being equally effective and used symmetrically and non-additively), while cueing with multiple elements would increase the chance of accessing a fragment containing a given element required for recall.

In a long-term study of his own memory, Wagenaar (1986) attempted to recall different autobiographical events recorded over four years, by probing himself with different

elements of each event (who, what, when and where). Consistent with Jones' (1976) study, he found that temporal information (when) was a very poor cue even though it could be retrieved reasonably well. However, in contrast to Jones he also found marked differences in the usefulness of the remaining elements as cues and asymmetry in their processing: He found 'what' to be the best cue, while 'where' was slightly better than 'who'. Correspondingly, 'what' was also used asymmetrically in being less well retrieved via other elements than they were retrieved by it. The observation that not all elements of an event will serve as equally effective retrieval cues is also stressed in the 'headed records' model of memory (Morton & Bekerian, 1986; Morton, Hammersley, & Bekerian, 1985). Finally, also unlike Jones, Wagenaar found that the advantage of retrieving an element by cueing with multiple other elements slightly *exceeded* that predicted by a fully independent model in many cases.

Wagenaar interprets the differences between his study and Jones' in terms of differences in cue specificity. In Jones' study, within each list of nine events, each cue was specific to only one event, whereas Wagenaar's cues varied in specificity, with 'what' the most specific, and who and where varying in specificity. Thus more specific cues might be more efficient in prompting retrieval, and many less-specific cues might, in Jones' terms, be contained in very many fragments. This latter consideration raises the possibility that multiple cues are combined into configural cues that could overcome the lack of specificity, as has also been suggested by Foss and Harwood's (1975) model of sentence recall. An alternative interpretation would simply be that the elements of Wagenaar's events were stored independently, perhaps differing from Jones' stimuli in being truly multi-modal. The slight increase in the advantage found for multiple cueing might result from the multiple-cue retrieval attempts occurring after the single-cue attempts.

Here we investigate the binding of the context of an event with the event's content in episodic memory using a computer based virtual reality (VR) paradigm involving pseudo-realistic simulated events. We hope to combine some of the contextual richness of autobiography with some of the control of stimuli across participants of traditional laboratory-based memory experiments. In this virtual context-dependent memory paradigm (or 'VCM' paradigm), participants move through the virtual environment and encounter virtual characters within it. The events for which memory will be tested

consist of the presentation of an object to the participant by a virtual character (see Figure 2). We distinguish between the 'content' of the event, i.e. the change in the world that marks the event – in this case the presentation of an object, and the ongoing 'context' of the event, including the surrounding spatial environment, time, and the person giving the object, see e.g. Burgess, Maguire, Spiers, O'Keefe (2001).

Participants experience a series of such events, as they move about the virtual town. After this learning phase participants are tested on their memory for the events using a context-dependent two-alternative forced choice paradigm: pairs of objects are presented in a particular place, with a particular character present. Different types of questions probe memories for different elements of the context of the events experienced in the learning phase (e.g. 'which object did you receive in this place?'). In addition, one question-type tests the memory for the content of the events alone (i.e. the object given), in which the familiar object that was present in the learning phase must be recognized compared to a similar-looking novel foil. In one variation of this paradigm, we added odor as an additional contextual element. Our scope was to explicitly look for any relationships between the probability of retrieving one element of an event and the probability of retrieving another element of that same event.

The categorization of memory into sub-types, such as 'episodic memory', goes hand-in-hand with consideration of its neural bases. Indeed the aim of much of the neuroscience research into memory is to match structure to function. As we discuss below, the hippocampus has been strongly implicated in supporting episodic memory. Here, we note that discussion as often concerns the specific role in memory played by the hippocampus as concerns the specification of the psychological process of 'episodic memory', exemplified by the term 'hippocampal-dependent memory'. Before we present the methods and results of our experiments in detail, we discuss the neuropsychological and neuroimaging findings relating to the neural bases of episodic memory, and of our VCM task in particular (see also Burgess, Maguire, & O'Keefe, 2002, for a review).

### Neural bases of episodic memory

There is a consensus that context-dependent memory for personally experienced events (i.e. 'episodic' memory) is supported by the hippocampus (Eichenbaum & Cohen, 2001, Kinsbourne & Wood, 1975, O'Keefe & Nadel, 1978, Squire & Zola-Morgan, 1991, Vargha-Khadem, Gadian, Watkins, Connelly, Van Paesschen, & Mishkin, M., 1997, for reviews see Spiers, Maguire, & Burgess, 2001c, Burgess, Maguire, & O'Keefe, 2002), albeit opinion remains divided about what possible other roles the human hippocampus might perform and over the role of other brain regions in episodic memory. This controversy includes debate concerning the possible hippocampal contribution to a-contextual forms of memory such as memory for factual knowledge ('semantic memory'). Declarative memory theory, for example Squire & Zola-Morgan (1991), sees the medial temporal lobe (including the amygdale and surrounding neocortex as well as the hippocampus) as supporting all forms of explicit memory, in an undifferentiated manner.

A further dissociation has been postulated between different types of retrieval: psychological studies (e.g. Mandler, 1980, 1991, Jacoby, Toth, & Yonelinas, 1993, see Yonelinas, 2002 for a review) suggest that two distinct processes are involved in recognition memory, one based on a general sense that the stimulus has been encountered before ('familiarity-based recognition'), the other entailing specific retrieval of an event and its context ('episodic recollection'). It has been suggested that these processes are dissociated in the brain, with a circuit including the mamillary bodies, anterior thalamus and hippocampus supporting episodic recollection, while a distinct parallel system, including the medial thalamus and perirhinal cortex supports familiarity-based recognition (Aggleton & Brown, 1999, Gaffan & Parker, 1996), Delay & Brion, 1969, see also Bogacz, Brown, & Giraud-Carrier, 2001, Holdstock, Mayes, Roberts, Cezayirli, Isaac, O'Reilly, & Norman, 2002, Tulving, 2001, Vargha-Khadem et al. 1997, Wan, Aggleton, & Brown, 1999, Yonelinas, 2002). A recent review (Rugg & Yonelinas, 2003) concluded that clinical data support the dual-process model, suggesting that, while familiarity is commonly impaired in amnesia, recollection is disrupted to a greater degree.

We have used the VCM paradigm introduced above in some recent neuropsychological investigations to address the issue of the neural bases of familiarity-based recognition and episodic recollection (King, Trinkler, Hartley, Vargha-Khadem, & Burgess, 2004, Spiers et al., 2001a, 2001b). First we discuss our experiments with Jon, a young man with focal bilateral hippocampal pathology (see Vargha-Khadem et al., 1997). He was between 5 and 6 years old, when it was discovered that he was experiencing spatial, temporal and episodic memory problems. Further investigation revealed selective bilateral hippocampal pathology apparently caused by perinatal anoxia. His hippocampal volumes are approximately half those of control participants (Gadian et al., 2000). There is also evidence that the remaining hippocampal tissue is compromised but that extra-hippocampal regions are largely preserved. Jon's educational record suggests few problems with semantic memory, for instance he obtained a GCSE in History. His verbal IQ was assessed to be 108 and his performance IQ 120 when tested at the age of 19.

We tested Jon using the VCM paradigm (for details see *Experiment 1* and Spiers et al., 2001a) and found that his ability to recognize the objects used in the events was spared but that his context-dependent recognition memory was impaired. However, there were two potential problems with this study. Firstly, control participants also showed lower scores for the context-dependent task than the object-recognition task in this experiment, so that we could not exclude the possibility that a non-linear effect of difficulty compromised Jon's performance. Secondly, we were concerned that a high degree of similarity between the contexts of different events might have reduced their distinctiveness, possibly introducing a lack of specificity in the contextual retrieval cues. These concerns were addressed in a second experiment (see *Experiment 2* and King et al., 2004) in which performance was matched across all conditions and each event occurred in a unique context. This experiment replicated the earlier findings, in that Jon was impaired in the context-dependent recognition tasks but not impaired when asked to recognize objects on the basis of their familiarity (see Table 1).

-----Table 1 about here-----

The pattern of performance shown by patient Jon conflicts with the declarative memory theory (Squire & Zola-Morgan, 1991) since both the spared item recognition and the

impaired context-dependent recognition are 'declarative' tasks. Note however that other patients with hippocampal damage have been described with impaired item recognition memory (Manns & Squire, 1999, see Spiers, Maguire, & Burgess, 2001c for a review). By contrast, the above findings do conform to the idea that the hippocampus and related structures along Papez's circuit (Papez, 1937) support episodic recollection while a separate circuit including perirhinal cortex support familiarity-based recognition (Aggleton & Brown, 1999, Gaffan & Parker, 1996, Baxter & Murray, 2001). Interestingly, the pattern of performance of patients who had had unilateral anterior temporal lobectomies (removing tissue from both hippocampus and perirhinal cortex) in the VCM paradigm was also consistent with the assumption of separate processes: Patients with *left* anterior temporal lobectomies were found to be significantly impaired on the context-dependent questions, while those with *right* anterior temporal lobectomies were found to be impaired on the object recognition question (Spiers et al., 2001b). The pattern shown by these patients and by Jon is consistent with hippocampal involvement in context-dependent memory (and more so on the left, consistent with functional neuroimaging data) and perirhinal involvement in object recognition (and more so on the right). See Burgess et al., (2002) for further discussion.

With regard to the nature of the episodic information stored by the medial temporal lobes, Marr's (1971) seminal hippocampo-cortical model of memory saw the hippocampus as providing a mechanism for the rapid storage of a simple representation of an event, from which semantic information could later be abstracted and stored in the neocortex. Importantly, these simple representations were thought to be formed of only those elements through which an event is later addressed, consistent with the ideas of cue specificity discussed with respect to Wagenaar's data. On the other hand, O'Keefe & Nadel's extension of the cognitive map theory to humans (1978, chapter 14, 380 ff.) is more similar to the holistic viewpoint of equal and symmetric cue processing. O'Keefe and Nadel take up Tulving's semantic-episodic memory distinction, opposing "memory for items independent of time or place of their occurrence" in the "taxon system" with "memory for items within a *spatio-temporal context*" in the hippocampal "locale system". Specifically, the locale system provides multiple channels of access for the retrieval of any of the relationships embodied in the map, such that any relationship in the map can be retrieved by activating any other portion of the map, whether or not these relationships were noticed at the time of input (see O'Keefe & Nadel, 1978,

p.384). Eichenbaum & Cohen's (1988; 2001) characterization of the hippocampus as supporting flexible-relational memory maintains the idea of flexibility from the cognitive map (e.g. information should be retrievable via a variety of cues), but puts more stress on pair-wise associations, and so need not necessarily imply holistic representation.

Another idea related to binding in episodic memory is that of the hippocampus as a 'convergence zone' (Alvarez & Squire, 1994; Murre, 1996; Moll & Miikkulainen, 1997; Damasio, 1989): linking information from different sensory modalities that are represented in disparate cortical areas. In Experiment 3, we added an olfactory component to our (visual) virtual reality events, in order to compare uni-modal and cross-modal binding. Olfactory cues have further been seen as particularly potent reminders of past experiences (which is sometimes referred to as the 'Proust phenomenon' after Proust's (1922) description of such an event, see e.g. Chu & Downes, 2000). Such a role might reflect the direct connections between primary olfactory regions and the hippocampus (Dade, Zatorre, & Jones-Gotman, 2002). Scientific evidence however rather undermined this notion (e.g. Bolger & Titchener, 1907, Davis, 1975, Herz, 1998, Rubin, Groth, & Goldsmith, 1984). The only evidence for privileged olfactory cueing of memory comes from Chu and Downes (2002) who found that solely odor cues enhance autobiographical memory retrieval in a second retrieval attempt (following a first memory search, cued by a label) compared to other cues. Alternatively, Rubin et al. (1984) argue that long term retention in olfactory memory is due to odor cues to memory suffering from less interference than verbal cues during the retention interval, a hypothesis that is supported by evidence for reduced retroactive interference in olfactory memory (Lawless and Engen 1977).

For the sake of completeness we should also mention that many areas outside of the medial temporal lobes also play important roles in episodic memory. For example, the frontal lobes are vital for the strategic organization of retrieval, editing, selecting, categorizing and inhibiting memories as appropriate (e.g. Burgess and Shallice, 1996). How the frontal lobes interact with the medial temporal lobes to provide a full episodic memory system is an area of increasing interest (see Wheeler, Stuss, & Tulving 1997, Simons and Spiers 2003 for reviews). The VCM paradigm described in this chapter has

also been used in functional imaging studies to investigate the wider neural systems supporting context-dependent memory. Memory for the spatial context (*where*) of an event was associated with hippocampal, parahippocampal, retrosplenial and medial and posterior parietal activations in two studies (Burgess, Maguire, Spiers, & O'Keefe, 2001, King et al. in prep). Moreover, these studies relate to the dual-process argument above, in that they did not find activation of the medial temporal structures for familiarity-based object recognition, consistent with our neuropsychological findings on the same tasks (Spiers et al., 2001a; King et al., 2004). They also relate to the role of the frontal lobes in episodic memory. In the first study (Burgess et al., 2001), the 16 events in a trial shared only two people and two places as their contexts (see Experiment 1 for details). The widespread lateral and anterior prefrontal activation found in this study, but not seen in previous studies of autobiographical memory (Maguire & Mummery, 1999, Maguire, Mummery, & Büchel, 2000) was interpreted as reflecting the interference resulting from such over-lapping contextual cues. This interpretation is consistent with neuropsychological studies (e.g. Incisa della Rocchetta & Milner, 1993) and with our more recent functional neuroimaging study, involving 20 events with distinct people and places (see Experiment 2) in which the prefrontal activations from the earlier study were much reduced (King et al., in prep.).

### ***Experiment 1***

In the first version of the virtual context-dependent experiment, participants experienced a series of 16 events that took place in two different places within a virtual reality environment, and with two different 'people' (virtual reality characters) from whom participants 'received' a different object at each encounter. This encoding phase was followed by a forced choice test of recognition memory probing each event 4 times addressing memory for different aspects. The whole sequence was then repeated with 16 new objects in 2 new places with 2 new characters. In total, participants answered 128 memory questions on the experienced VR events. (For further details see Spiers 2002).

## Methods

### Participants

35 participants (9 female, 26 male) took part in the study, with an age-range of 18 to 33 years (mean 25 years). Their mean IQ was 105 (inferred from mean score on Ravens Progressive Matrices, Set 1,  $M = 9.5$ ,  $SD = 1.7$ ).

### Encoding Task

Participants followed a marked route through a virtual reality town (designed using Duke Nukem 3D, see Burgess et al., 2001). They repeatedly encountered one of two characters in one of two rooms along the route (not always in the same part of the room). When they encountered a character, they pressed a key, causing the virtual character to present an object (e.g. a light bulb). Participants were told that they would subsequently be tested on which of two objects they had received, who gave them each object, where they received each object and in which order they received them.

### Recognition Test

Immediately after each encoding phase, participants performed 4x16 forced choice recognition trials. For each they re-entered one of the two rooms (in a counterbalanced sequence), encountered one of the two characters, and saw two objects that appeared on the nearest wall along with a word indicating the type of memory question. There were four different memory questions as follows: 'Object' - "Which of the two objects displayed were you given?"; 'Person' - "Which of the two objects did you receive from the present character?"; 'Place' - "Which of the two objects did you receive in this location?"; 'First' - "Which of the two objects did you collect first?". For the 'Object' question, the foil was a similar looking version of the original, while foil objects in the other conditions were from other events in the encoding task, thus equally familiar.

### Practice trial

Participants were given a practice trial during which they followed each of the two routes, encountered two characters who presented them with four objects in different places. They were then given one of each type of question concerning memory for these objects and asked whether they had used any particular encoding strategies. If they did, they were asked to refrain from using these strategies in the test and to simply pay attention to the various elements of each event.

## Results

There was a significant effect of question type on performance (see Figure 1). Post-hoc single comparisons revealed significant differences for each pair ( $t(34) > 3.5$ ,  $p < 0.01$ ) except for the Person versus First comparison. In particular, object recognition (the Object task) was most successful. Of interest regarding the question of binding is the comparison between the context-dependent question types (Person versus Place versus First). Overall, retrieval was more successful via the Person cue and temporal cue First than the Place cue.

----- Figure 1 about here -----

We then wished to find out whether retrieving the object associated with one element of an event was correlated with retrieving that object via another element of the *same* event. We present *theoretical* contingency tables for a pair of question types under a Fully Dependent Model (where retrieving one element of an event is maximally correlated to retrieving another element of the same event) in Table 2, and under a Fully Independent Model (where retrieving one element of an event is independent of retrieving another element of that event) in Table 3. In the Fully Dependent Model, if one retrieval cue is more successful than another, then both cues should be successful every time the least successful is, hence cell **a** (proportion correct for both questions) represents the proportion correct of the least successful ( $a = p$ ). Also, there should be no cases where the least successful is correct and the more successful incorrect, hence cell **b** = 0. Further, all incorrect cases of the more successful cue ( $1-f$ ) must occur in the case where neither cue is successful, cell **d** =  $1-f$  and finally, cell **c** expresses the case where only the more successful cue retrieves the correct answer ( $c = f-p$ ).

----- Table 2 about here -----

In the Independent Model, the proportions of **a**, **b**, **c** and **d** can be estimated by combining the assumed probabilities of correct and incorrect cases of two cues, see Table3.

----- Table 3 about here -----

Observed responses and predicted values from the Fully Dependent and Independent Models were subjected to a Chi squared analysis, Table 4 shows the corresponding statistics and p-values.

----- Table 4 about here -----

Four out of the six contingency tables (i.e. question pairs) showed statistically significant differences between the Dependent Model and the observed data, whereas there was a close fit between the data and the Independent Model in all cases, suggesting that the associations formed between objects and context are encoded independently in memory.

Additionally, each individual participant's data was analyzed using a corrected correlation statistic as suggested by Hayman & Tulving (1989), in order to avoid 'Simpson's paradox' when interpreting contingency tables from group data. Results at the single participant's level corroborate the above findings on group data (for details see Spiers 2002).

### Discussion

Evidence from Experiment 1 casts doubt on the suggestion that events are encoded holistically in our hippocampal-dependent episodic memory test. A very good fit of the data is provided by a model assuming an independent probability of success in retrieving the same event from different contextual elements. The independence between performance in the object-question and the context-dependent questions might be specifically related to use of an additional process of familiarity-based recognition that does not depend on the hippocampus (see King et al., 2004). Further, performance in the 'First' question might also be influenced by an additional factor as this question requires retrieving and comparing the place in temporal sequence of both objects (with the exception of the first and last objects received). However, a holistic encoding theory would at the very least predict some dependence between performance on the place and person questions, and no such dependence was observed.

The generally low performance in the context-dependent questions is of concern because it might imply a high degree of guessing. Random answers could therefore be obscuring some possible dependencies in the data. Furthermore, the high degree of interference between 16 events involving only two characters and two places may have resulted in recollection being less 'truly episodic' in Tulving's sense of fully re-experiencing distinct events. Similarly, the re-use of contextual cues in different events might prevent simple use of a fragmentation model such as that in Jones (1976), as discussed by Wagenaar (1986). Finally, performance differed between the various question types which in itself rules out complete dependence between performance in different questions relating to a given event. These issues were addressed in Experiment 2.

## ***Experiment 2***

In this experiment we used a VCM paradigm involving 20 events with unique contexts, each involving a distinct virtual character and location. We again attempted to look at whether the probability to retrieve an event via one contextual cue was dependent or independent of the probability to retrieve the same event via another contextual cue. In addition we attempted to equalize performance across the question types. For further details see King et al. (2004).

### *Methods*

#### *Participants*

12 male participants (who were age- and IQ-matched for comparisons of performance with patient Jon, see Introduction) took part in this experiment, age range = 21-28, mean age 23.4, mean IQ 114 (inferred from Ravens Advanced Matrices, Set I, mean score 10.43, SD 1.22).

#### *Encoding Task*

A virtual reality town, built on the commercially available Computer Game Deus Ex, provided the environment for the test. It was presented on an AMD Athlon XP2200 computer with a standard 19 inch monitor at a resolution of 800x600 pixels and a vertical refresh rate of 60Hz. To manoeuvre within the town, participants used the

cursor keys of the keyboard and followed a trail of green icons, see Figure 2a. In distinct places along the route, participants encountered virtual characters who presented them with an image of an object (display size 7x7 cm), see Figure 2. Subsequently, a new trail of icons would appear for the participant to follow to the next encounter. Participants were told that they would be tested on these events afterwards and instructed to try and remember the person, object and place of each event. All participants experienced the same sequence of 20 events (rather than counterbalancing order, objects etc.) as the data were also used to assess the memory performance of patient Jon. The encoding phase took about 15 minutes on average.

----- Figure 2 a and b about here -----

### Recognition Task

Immediately after the encoding task, participants were given paired forced-choice recognition tests on all aspects of all events (3x20 tests): they were presented with two objects, on the left and right of the screen, a virtual character in the foreground and a snapshot of one of the locations in the background. A word appearing on the top of the screen indicated what type of event information was being probed (see Figure 3).

----- Figure 3 about here -----

Two questions probed context-dependent memory ('Place' and 'Person') and one question ('Object') probed recognition of the content of the events, as in Experiment 1. Participants responded by button press, indicating whether the left or the right object was associated with the cue in question. In the Object condition, the foil object was a similar looking version of the original image that had been presented in the encoding task.

### Practice Trial

Prior to testing, participants were given a trial run of both the encoding task (consisting of 3 events presented in an alternative virtual reality town) and the recognition memory test.

## Results

The average performance over all participants is shown in Table 1. Note that average performance does not differ between the different question types. As before, our main focus of interest was whether performance on the different question types was correlated across events or not. As in Experiment 1, we constructed contingency tables for each pair of questions (e.g. Person and Place) for each participant. This time we analyzed the contingency table for each participant individually using Fisher's exact test. There was a good match between the observed results in that the Independent Model assumed by Fishers exact test was far from being rejected, see Table 5.

----- Table 5 about here -----

We also explicitly created the contingency tables expected under an Independent Model (as explained in Experiment 1, see Table 3), on basis of the frequencies of each type of paired response (e.g. correct-correct, correct-incorrect etc.) for the 20 events. Because performance was approximately equal across conditions in Experiment 2 we were also able to create a Fully Dependent model that included guessing (see Table 6).

----- Table 6 about here -----

The Chi Squared Test applied to a contingency table over all participants corroborates data from the single participants' analysis, see Table 7 – fitting the Independent Model and rejecting the Dependent Model With Guessing.

----- Table 7 about here -----

Furthermore, we sought to directly compare the two models, evaluating the difference between both models and the data, see Table 7b. For all three pairs of questions compared, the sum of squared differences between model and data is significantly smaller for the Independent Model than for the Dependent Model With Guessing.

In summary, as in Experiment 1, we found no evidence in favor of events being encoded holistically.

### Discussion

The results of Experiment 2 further support a model where all elements of an event are encoded and retrieved as independent pair-wise associations. In this experiment some of the concerns regarding the interpretation of Experiment 1 could be eliminated. Performance was reasonably high, suggesting a reduced role for guessing, and also well-matched across conditions, removing one potential obstacle to finding evidence in favor of a fully dependent model.

There is one remaining difficulty in interpreting the results from Experiment 1 and Experiment 2 in terms of whether or not retrieval is holistic i.e. all-or-none: Even if participants answer the Person-question correctly but not the Place-question regarding the same event, there remains the possibility that in the instant of retrieving the Person information, they successfully retrieve the whole event, but that at the instant of retrieving the Place information, they fail to recall any elements of the event. Thus, our simple cued-recognition paradigm only allows us to conclude that events are not *encoded* holistically – in which case variations in the strength of encoding of different events would produce some dependencies among the performance of the different questions concerning the same event. However, it is still possible that events are *retrieved* holistically, because a separate retrieval process is required for each question regarding a given event. To address this issue, in Experiment 3 we added a cued-recall test to our paradigm, in which memory for different elements of an event could be probed simultaneously.

### **Experiment 3**

In this experiment we made a further modification to the VCM paradigm to test cued recall in addition to forced choice recognition of context-object pairs and familiar objects. After the encoding phase, we additionally presented participants with the individual components of all events and asked them to reconstruct the events they had

experienced. This also allowed us to look at possible "retrieval-hierarchies", i.e. whether some contextual cues may be preferred over others to retrieve information about an event. Furthermore, in this experiment, each VR-event included a distinct olfactory cue in addition to distinct people and places. This allowed us to begin to investigate retrieval of events via truly cross-modal contextual cues.

## Methods

### Participants

12 participants participated in the experiment, 5 females and 7 males, with an average age of 28 years (ranging from 22-39), mean IQ = 108 (inferred from Raven's Advanced Matrices, Set 1, average 9.75, SD = 2.0). Only people who rated their sense of smell 5 or above were included (self-rating, on a scale from 1 to 7).

### Practice Trial

Prior to testing, participants were given the same trial run as in Experiment 2 with the addition of an odor cue for each event.

### Pre-experimental Exposure to Olfactory Stimuli

After participants had conducted the practice trial, they were presented with the 10 phials used in the experiment, and asked to sample all smells, one after another, and to verbally describe them. If participants failed to come up with a label, they were given a hint (e.g. "is it a flower?") and further prompted until they had a specific label for each of the ten odor stimuli.

### Encoding Task

The virtual town, computer and manner of navigating and responding were identical to Experiment 2. In distinct places along the route, participants would meet virtual characters who would present them with an object. Simultaneously with the occurrence of the object, the experimenter would also present an odorous stimulus from a phial (about 1cm from the participant's nose) for the duration of one sniff. Then, participants would continue their journey through the virtual town along a new trail of green icons. Participants were told that they would be tested on these events subsequently and instructed to try to remember the person, object, place and odor of each event.

Each participant experienced a unique composition of 10 events, that is, assignment of objects, people, places and odors was randomly varied between participants. There were 5 different possible first places that were always reached from the same start location within the town and the sequence of locations was the same for all participants. The whole "encoding-walk" took participants about 15 minutes on average.

### Odor Stimuli

Odors were presented in medicine bottles labeled with numbers visible to the experimenter only. Through extensive pilot experiments we sought out odorous liquids that were rated neutral in hedonic quality, were matched in perceived intensity, and easily distinguished from one another. They were rated 'familiar' and could be described verbally, e.g. 'rose', 'peanut butter', 'white spirit', 'spearmint', etc.

### Recognition Test

After the encoding task, participants were given paired forced-choice recognition tests: A question word on the screen preceded the presentation of two objects on the left and right of the screen, together with one contextual cue corresponding to the question word: In the Place condition, the two objects appeared in front of a snapshot of one of the 10 events-locations, in the Person condition, one of the 10 characters appeared between the two objects in front of a plain brown background, in the Odor condition, the participants were presented with one of the 10 odors from a phial together with the visual presentation of two objects in front of a plain brown background. Participants indicated by button press, whether they had received the left or the right object in the presence of the respective cue. For the Object condition, an object from the encoding phase was presented together with a similar looking new lure in front of a plain brown background.

### Cued Recall Test

After the recognition memory task, participants were shown randomly arranged laminated paper-copies of all elements of the events they had experienced in the encoding task: virtual characters, images of objects and snapshots of locations. The odor stimuli were presented on commercially used test-strips on pegs. Participants were

instructed to try and reconstruct as well as possible "what they could remember". They were further told that their reconstruction process would be recorded on-line by the experimenter. We recorded which card / odor-clip was put together with which other card / clip and in what sequence. Participants were allowed to finish before they had recombined all single elements, when they felt they "could not remember anything more". Finally, participants were asked what strategies they had used to encode the events.

## Results

### Recognition Test

Table 8 shows the results of the recognition memory test in Experiment 3.

----- Table 8 about here -----

Performance on the object question was significantly greater than performance on any other question type (repeated measures-ANOVA, over all  $F=15.7$ ,  $df=9$ ,  $p<0.001$ , simple contrasts between object and any other condition,  $F$ 's= 51.5, 24 and 20.4 respectively,  $df=1$ , all  $p<0.001$ ) and better than in Experiments 1 and 2, probably because of the smaller number of events used. Within the context-dependent question types, we found significantly better performance for Person questions than for either Place or Odor questions (paired t-tests, one-tailed,  $p<0.01$  for person vs odor,  $p<0.05$  for person vs place), but no significant difference between performance on Place and Odor questions.

Regarding the odor cue, we assessed a possible relation between recognition memory score and odor identification ability: Odor identification as assessed pre-experimentally was estimated 1 for a correct label, 0.5 for an approximate description and 0 for no description. There was no correlation between individual *participants'* identification scores and their memory scores for odor-cued recognition ( $R^2=0.0048$ ), however, there was a high correlation between an *odor's* identification score and how well it elicited odor-cued recognition ( $R^2=0.80$ ).

### Cued Recall

Data from participants' reconstruction processes were scored as follows: As participants were allowed to continue reconstructing one event after having attempted to reconstruct another event in the meantime, and made frequent use of this possibility, all participants' reconstructions of all 10 events were scored in two ways. Firstly, we counted how many elements of an event a participant correctly put together in the first attempt, before going on to reconstructing another event. We refer to this as the '*initially remembered*' score. Secondly, we counted how many elements of an event participants had correctly put together at the end of their reconstruction process. We refer to this as the '*eventually remembered*' score.

Overall, for '*initially remembered*' items, 66% of all events (total = 10 events x 12 participants) were at least partially correctly reconstructed as opposed to 34% of events for which no two elements were correctly matched. For '*eventually remembered*' items, these percentages amounted to 81% at least partially correctly reconstructed events versus 19% entirely forgotten (see Table 9).

----- Table 9 about here -----

The percentage of events for which all 4 elements were remembered together correctly does not exceed 19%. The full retrieval of a complete event is thus rare compared to the retrieval of the object via a single cue in the recognition test (which was 64% for the worst case, odor). Table 10 shows which elements were not recombined with the others correctly when the other 3 elements were correctly matched, revealing that the association with odor was the weakest over all: it could not be reattributed to the other elements of an event in 91% of all cases that missed one element. Note that initial and eventual reconstruction cannot be compared directly, as some events may have become fully reconstructed in the meantime and are thus no longer counted in the 3-elements-category, or participants may have taken the element-combinations apart again and recombined them differently.

----- Table 10 about here -----

Events were retrieved via the following cues first: person 38%, object 18%, odor 7% place 5% (note that these percentages do not sum up to the total number of 66% 'initially remembered' because some events were retrieved by 2 cues simultaneously). Participants thus show an overall cue preference in favor of the Person cue over the Place and Odor cues in our cued-recall setting, where the cues themselves do not have to be recalled from memory but are already present and merely have to be combined correctly.

### Strategies

Participants did not seem to employ distinct strategies throughout. However, they often gave examples of well remembered associations, which hinted towards a facilitation for associations with a common *semantic* theme, that was either inherent, e.g. “chef” and “kitchen” or easy enough to be thought of e.g. "cream-colored glove and smell of coffee-liquor". Note that all aspect-combinations per event, and thus their semantic coherence varied between participants.

### Discussion

The cued recall experiment revealed clearly that participants did not successfully retrieve complete events consisting of all four testable elements, neither at an initial retrieval attempt nor after the final retrieval attempt. On the other hand, the percentage of events for which more than 2 elements are correctly associated together increases from 25% initially remembered to 48% eventually remembered. Participants are thus often successful at re-attempting retrieval after a first attempt, and mostly add a third (and eventually fourth) element. This could reflect the use of (independent) pair-wise associations with either of the elements of the first pair.

In contrast to Experiment 2, which shared the virtual reality setting of Experiment 3 (but not the odor-element), performance in the Object condition was much higher than in the other (context-dependent) conditions of the recognition test. Possible explanations are that: firstly, the total number of events was only 10 as compared to 20; secondly, the number of repeated exposures to objects in the recognition test was increased because of an additional Odor condition; and thirdly, that presentation of the

odor distracted participants from the other elements of context. These last two factors may have also contributed to the worse performance in the context-dependent question types in this experiment. We note however that we previously found that retrieval of different elements of an event was independent whether difficulty was matched (Experiment 2) or not (Experiments 1).

Retrieval *independence* is complemented by retrieval *asymmetry*: The Person cue featured predominately as the element by which other elements of an event were retrieved. Odor, by contrast, was most noticeably forgotten and hardly served as a primary cue for the reconstruction of the event. Similarly, performance of the odor cue in the forced-choice recognition test was worst. The success of a particular odor cue stimulus was correlated with how nameable it was considered to be over all, thus indicating odor memory facilitation through available semantic information, as also shown by Rabin and Cain (1984). However, this correlation was not found within-participants, suggesting that performance was not necessarily influenced by a participant's (momentary) ability in naming (that could in turn be due to a "tip-of-the-nose-state", Lawless and Engen 1996), but rather the relative nameability of the stimuli in general.

## ***General Discussion***

We used a virtual reality paradigm to study memory for the context and content of a series of pseudo-realistic events in which the participant encounters a person in a place and receives an object from them. Memory for context was tested via paired forced choice of which of two objects were received in a given place or from a given person (context-dependent memory). Memory for the content of an event took the form of a paired forced choice of which of two similar looking objects had been received before (object recognition). In neuropsychological and functional neuroimaging experiments we found the hippocampus to be implicated in the context-dependent memory questions, but not in the object-recognition questions. We also correlated memory performance for each contextual aspect and for the content per event, in order to test whether encoding of these events was holistic or fragmented.

### Events are not encoded as holistic units

We found that participants' retrieval success when cued with one element of an event does not correlate with retrieval success when cued with another element of that same event (Experiment 1). Moreover, this finding holds when the overall retrieval performance is the same on average for retrieval cued by the different elements, and when each event occurs in a distinct context (Experiment 2). We thus conclude that, in our experiments, events are not *encoded* holistically since this would predict dependencies between the retrieval of the same event by different cues. By contrast, a model based on independent pair-wise associations between elements provides a good fit to the data.

Whether or not each *retrieval* of an event might be holistic (i.e. all-or-none), was investigated by using an additional cued-recall test, in which the participants had to recombine all of the individual elements into the events that they had experienced (Experiment 3). We found that at the initial retrieval attempt only 7% of all events were retrieved fully, with 59% remembered partially and 33% not remembered at all. Importantly, this performance increases to 19% entirely remembered in the end, as opposed to 62% partially remembered and 19% forgotten completely. This suggests that recollection of events is also not holistic, but rather is partial and iterative in nature: more and more information is added in subsequent retrieval attempts.

### Relation to previous work

Our results are inconsistent with the idea that episodic recollection corresponds to 're-experiencing' an event complete with its multi-modal context in such a realistic way that a mechanism of auto-noetic awareness (e.g. Tulving, 2002) is required to disambiguate it from current perception. They are also inconsistent with the spirit of the 'locale-system' proposed by O'Keefe & Nadel (1978) in which an event is stored in a map-like set of relations such that it can be equally well retrieved via different relationships.

Is it possible that our stimuli somehow fail to capture the essence of autobiographical episodic information? For example, even though we took care to use distinct contextual *elements* for each event in Experiment 2, all of the events involved the same *action*: walking up to a person and 'receiving' the object that appears as a result of that encounter. In Brewer and Dupree's (1983) study, different goal-directed actions, viewed on film, were remembered holistically. It may be that the similarity of the actions in our events caused interference between them that disrupts the holistic and distinct recollection of each one. However, this interpretation is undermined by the similarity between our results and those in Wagenaar's (1986) study of his own autobiographical memory. Wagenaar found that some elements of an event formed better cues than others and also found that multiple cueing by different elements of an event increased the probability of retrieval in line with (and sometimes exceeding) the prediction of independent pair-wise associations. In addition, he reports many events that were only partially remembered, and a failure to retrieve around 20% of events (consistent with the final result in the cued recall of our Experiment 3).

Our results also contrast with Jones' (1976) finding of independent but holistically encoded fragments. Or, put another way, we only found evidence for fragments including pairs of elements but not triples. As argued by Wagenaar (1986) some differences might be due to Jones' paradigm, using nine lists each of nine objects in nine different colors and nine different locations. In this paradigm, participants might quickly learn that, within a list, each element of an event is unique and thus any fragment containing a given element will be specific to the single event containing that element. In autobiographical studies, and in our VCM paradigm, the participant will not in general be able to make that assumption (even though it would have been correct in Experiments 2 and 3, the participant would not have had time to learn that – getting only one full trial). In the case of multiple fragments potentially containing the same single elements a simple fragment model may not be sufficient to describe performance, and other processes may become the performance-limiting factor. For instance, as Wichawut & Martin (1971) found, retrieval independence is related to the strength of a formed association. They found that, after learning A-b and A-C associations, the responses B and C are retrieved independently so long as at least one of the pairs is well-stored in memory, but that retrieval dependencies arise if both are weakly stored.

### Retrieval cue hierarchy

We found that different levels of access to the memory of the content of the event (the object) are afforded by different contextual cues. This is consistent with Wagenaar's findings. It is also consistent with Marr's (1971) model, the 'filing cabinet' model referred to by Wagenaar (1986), and the model of headed records (Morton & Bekerian, 1986, Morton et al., 1985) in which some elements of a memory are seen to be efficient retrieval cues (e.g. the name of a person) but are much less easily retrieved themselves.

In our cued recall experiment, the Person cue was most frequently chosen to start retrieval of episodic information. At first sight, this contrasts with Jones' (1976) findings that cued recall was symmetric in that the probability for either of two (perceptual) components of an event to promptly recall the other was the same. There are several potential explanations to account for this. As well as the chances for participants to evaluate cue-specificity over several trials (discussed above), the elements of context in Jones' experiments perhaps had more similar, and reduced, semantic complexity than the elements of context in our experiment. In Jones' experiment, one element ('location') was one of 9 grid positions on a backdrop and was thus similar to the other element ('color') in its (lack of) semantic complexity. Asch & Ebenholtz (1962) similarly demonstrate approximate symmetry in the recall of two-component visual patterns but argue that asymmetry in other circumstances could be due to differential availability of the components, perhaps due to differential levels of semantic association.

Clayton et al. (2001) suggest that 'where' is the predominant element binding an episode together in their investigation of episodic-like memory in scrub jays compared to 'what' and 'when'. This clearly contrasts with our findings from Experiment 1 where recognition performance was equal for Person and First (the temporal order question) and better than for the Place question. We hypothesize however that retrieval cue success (and preference) depends on the circumstances: the crucial cue of retrieval might well shift away from Place to People, depending on the nature of information to be remembered. Whereas, e.g. for scrub jays caching food, the most successful memory

cues might well be places (triggering additional episodic information), for human participants wandering around in (virtual reality) towns, the most relevant cues would be the people who provide them with objects. The Person might be given preference because she could hypothetically walk away and disrupt the 'where' whilst taking the 'what' away with her. Clayton et al.'s original paradigm (Clayton & Dickinson, 1998, Clayton, Griffiths, Emery, & Dickinson, 2001) did not test memory for 'who' was involved in a caching event, but see Emery & Clayton (2001). An alternative explanation would be that cue preference is dependent on the distinctiveness of cues of the same category across events. If for example the places are very similar, other cues will contribute more to the distinctiveness of the event.

### *Role of semantics in episodic memory*

In our experiment, where we had created episodes of random semantic consistency, there were only a few combinations of places, people, objects and smells, created by chance, that were inherently consistent (imagine for example 'kitchen', 'chef', 'cup and saucer' and 'smell of peanut-butter' in Experiment 3). Notably, participants happily made use of any semantic consistency (which they reported in the post-experimental assessment of strategies used). Furthermore, we found that those odorants which were more easily given a label (in a perceptual test before the VCM experiment) and thus a meaning, proved to be better cues to episodic memory. Thus, semantic relations may play a role in binding of episodic memory, as emphasized by Tulving and Markowitsch (1998; "Encoding of information into the episodic system depends critically on the semantic system", p.200), and as shown in earlier experimental investigations of human memory (e.g. Jenkins & Russell, 1952; Deese, 1959) and recently in people with dementia (Rusted et al., 2000). However, at the same time, our example also highlights the problem of knowing which of several semantically consistent alternatives to use. For example, how does one succeed in recalling that the 'chef' was paired with the 'cup and saucer' and not the 'teapot'? In a future experiment, we could test whether semantically related items tend to be confounded (see also 'false memory phenomena', Roediger & McDermott, 1995) and most interestingly, whether this would affect single pairings within an event as would be predicted by an independent model of encoding and retrieval, or whether its affect would be more holistic.

### *Olfactory cues are not especially evocative*

Despite the use of familiar, distinct and identifiable odors, the success of olfactory cues to retrieve episodes was relatively small compared to the other cues (place and person). Taking into account that perceptual features might account for this, and admitting that there is no information available about the extent to which the cues were matched semantically and perceptually across modality, we would nonetheless like to add another potential explanation. Olfactory stimuli are generally not easily tagged with a label, despite their perceived familiarity and the fact that the very same stimuli had just been labeled some minutes ago (e.g. Engen & Ross, 1973, Cain & Potts 1996), with participants sometimes becoming trapped in a "tip-of-the-nose-state" (Lawless & Engen, 1977). The binding of label and olfactory percept is volatile. So, on the one hand, semantic integration of an odor enhances its success as a retrieval cue, as reported above and shown previously (Lawless & Engen, 1977, Rabin & Cain, 1984, Lyman & McDaniel, 1990). But on the other hand an olfactory cue's frequent temporary failure to elicit a label might result in it being preserved in memory as a rather isolated and inaccessible trace. As such it might be a poor contextual cue, although by being recalled relatively rarely it might also remain a highly distinctive cue. Corroborating this 'rarity-argument', there is indeed experimental evidence showing reduced retroactive interference in olfactory memory compared to other modalities (see Lawless & Engen, 1977, Rubin et al., 1984).

### ***Conclusion***

In a chain of pseudo-realistic events which consist of the same element-categories throughout, each event appears to be encoded in terms of independent pair-wise associations between its elements. We found no evidence of a more holistic encoding in which all elements are associated together. On the contrary, performance on remembering the content of an event via one element of its context appeared to be independent of performance in remembering it via a second element. This finding argues against the idea that whole events are the units of episodic memory and are necessarily re-experienced in all their detail at retrieval.

The context-dependent memory task was shown to be hippocampal-dependent, in accordance with the idea that episodic memory is supported by the hippocampus (Eichenbaum and Cohen, 2001, Kinsbourne and Wood, 1975, Mishkin et al., 1997, O'Keefe and Nadel, 1978, Squire and Zola-Morgan, 1991). Our results thus also argue that the hippocampal contribution to episodic memory is not specifically to encode events in a 'map' of associations in which all elements can contribute equally. Elements of an event can be retrieved individually and in various combinations and, in our experiment of cued recollection, most favorably via the Person cue as opposed to the Place, Object or Odor cues. Finally, in our experiment including olfactory stimuli, we found that olfactory information is less closely tied to the event and serves as a less potent retrieval cue than the other elements of our events. This might be linked to poor semantic representation of olfactory information.

## References

- Aggleton, J. P. & Brown, M. W. (1999). Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *The Behavioral and Brain Sciences*, 22, 425-490.
- Alvarez, P. & Squire, L. R. (1994). Memory consolidation and the medial temporal lobe: a simple network model. *Proceedings of the National Academy of Sciences of the United States of America*, 91, 7041-7045.
- Asch, S. E. & Ebenholtz, S. M. (1962). The principle of associative symmetry. *Proceedings of the American Philosophical Society*, 106, 135-163.
- Baxter, M. G. & Murray, E. A. (2001). Opposite relationship of hippocampal and rhinal cortex damage to delayed nonmatching-to-sample deficits in monkeys. *Hippocampus*, 11, 61-71.
- Bogacz, R., Brown, M. W., & Giraud-Carrier, C. (2001). Model of familiarity discrimination in the perirhinal cortex. *Journal of Computational Neuroscience*, 10, 5-23.
- Bolger, E. M. & Titchener, E. B. (1907). Some experiments on the association power of smells. *American Journal of Psychology*, 18, 326-327.
- Brewer, W. F. & Dupree, D. A. (1983). Use of plan schemata in the recall and recognition of goal-directed actions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 117-129.
- Burgess, N., Maguire, E. A., Spiers, H. J. & O'Keefe, J. (2001). A temporoparietal and prefrontal network for retrieving the spatial context of lifelike events. *Neuroimage*, 14, 439-453.
- Burgess, N., Maguire, E., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, 35, 625-641.
- Burgess, P. W. & Shallice, T. (1996). Confabulation and the control of recollection. *Memory* 4(4), 359-411.
- Cain, W. S. & Potts, B. C. (1996). Switch and bait: Probing the discriminative basis of odor identification via recognition memory. *Chemical Senses*, 21, 35-44.
- Chu, S. & Downes, J. J. (2000). Odor-evoked autobiographical memories: psychological investigations of proustian phenomena. *Chemical Senses*, 25, 111-116.
- Chu, S. & Downes, J. J. (2002). Proust nose best: odors are better cues of autobiographical memory. *Memory and Cognition*, 30, 511-518.
- Clayton, N. S. & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, 395, 272-274.
- Clayton, N. S., Griffiths, D. P., Emery, N. J., & Dickinson, A. (2001). Elements of episodic-like memory in animals. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 356, 1483-1491.
- Dade, L. A., Zatorre, R. J., & Jones-Gotman, M. (2002). Olfactory learning: convergent findings from lesion and brain imaging studies in humans. *Brain*, 125, 86-101.

- Damasio, A. R. (1989). The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation*, 1, 123-132.
- Davis, R. G. (1975). Acquisition of verbal associations to olfactory stimuli of varying familiarity and to abstract visual stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1, 134-142.
- Deese, J. (1959). Influence of inter-item associative strength upon immediate free recall. *Psychological Reports*, 5, 305-312.
- Delay, J. & Brion, S. (1969). *Le syndrome de Korsakoff*. Paris: Masson.
- Duzel, E., Yonelinas, A. P., Mangun, G. R., Heinze, H. J., & Tulving, E. (1997). Event-related brain potential correlates of two states of conscious awareness in memory. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 5973-5978.
- Eichenbaum, H. & Cohen, N. J. (1988). Representation in the hippocampus: what do hippocampal neurons code? *Trends in Neurosciences*, 11, 244-248.
- Eichenbaum, H. & Cohen, N. J. (2001). *From Conditioning to Conscious Recollection : Memory Systems of the Brain*. Oxford: Oxford University Press.
- Emery, N. J. & Clayton, N. S. (2001). Effects of experience and social context on prospective caching strategies by scrub jays. *Nature*, 414, 443-446.
- Engen, T. & Ross, B. M. (1973). Long-term memory of odors with and without verbal description. *Journal of Environmental Psychology*, 100, 221-227.
- Fisher, R. P. & Chandler, C. C. (1991). Independence between recalling interevent relations and specific events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 722-733.
- Foss, D. J. & Harwood, D. A. (1975). Memory for sentences: Implications for human associative memory. *Journal of Verbal Learning and Verbal Behavior*, 14, 1-16.
- Gadian, D. G., Aicardi, J., Watkins, K. E., Porter, D. A., Mishkin, M., & Vargha-Khadem, F. (2000). Developmental amnesia associated with early hypoxic-ischaemic injury. *Brain*, 123, 499-507.
- Gaffan, D. & Parker, A. (1996). Interaction of perirhinal cortex with the fornix-fimbria: memory for objects and "object-in-place" memory. *The Journal of Neuroscience*, 16, 5864-5869.
- Hayman, C. A. & Tulving, E. (1989). Contingent dissociation between recognition and fragment completion: the method of triangulation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 228-240.
- Herz, R. S. (1998). Are odors the best cues to memory? *Annals of the New York Academy of Sciences*, 855, 670-674.
- Holdstock, J. S., Mayes, A. R., Roberts, N., Cezayirli, E., Isaac, C. L., O'Reilly, R. C., & Norman, K. A. (2002). Under what conditions is recognition spared relative to recall after selective hippocampal damage in humans? *Hippocampus*, 12, 341-351.
- Incisa della Rocchetta, A. & Milner, B. (1993). Strategic search and retrieval inhibition: the role of the frontal lobes. *Neuropsychologia*, 31, 503-524.
- Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General*, 122, 139-154.

- Jenkins, J. J. & Russell, W. A. (1952). Associative clustering during recall. *Journal of Abnormal and Social Psychology*, 47, 818-821.
- Jones, G. V. (1976). A fragmentation hypothesis of memory: cued recall of pictures and of sequential position. *Journal of Experimental Psychology: General*, 105, 277-293.
- King, J. A., Trinkler, I., Hartley, T., Vargha-Khadem, F., & Burgess, N. (2004). The hippocampal role in spatial memory and the familiarity-recollection distinction: a single case study. *Neuropsychology*, 18, 405-417.
- King, J. A., Hartley, T., Spiers, H.J., Maguire, E. A. & Burgess, N. Anterior prefrontal involvement in episodic retrieval is caused by contextual interference. In prep.
- Kinsbourne, M. & Wood, F. (1975). In Deutsch D. & Deutsch, J. A., (Eds), *Short-Term Memory* (pp 257-291). New York: Academic Press.
- Lawless, H. T. & Engen, T. (1977). Associations to odors: Interference, memories, and verbal labelling. *Journal of Experimental Psychology: Human Learning and Memory*, 3, 52-59.
- Lyman, B. J. & McDaniel, M. A. (1990). Memory for odors and odor names: Modalities of elaboration and imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 656-664.
- Maguire, E. A. & Mummery, C. J. (1999). Differential modulation of a common memory retrieval network revealed by positron emission tomography. *Hippocampus*, 9, 54-61.
- Maguire, E. A., Mummery, C. J., & Büchel, C. (2000). Patterns of hippocampal-cortical interaction dissociate temporal lobe memory subsystems. *Hippocampus*, 10, 475-482.
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, 87, 252-271.
- Mandler, G. (1991). Your face looks familiar but I can't remember your name: A review of dual process theory. In E. William, E. Hockley, & E. S. Lewandowsky (Eds.), *Relating Theory and Data: Essays on Human Memory in Honor of Bennet B. Murdock* (pp. 207-225). Hillsdale NJ: Erlbaum.
- Manns, J. R. & Squire, L. R. (1999). Impaired recognition memory on the Doors and People Test after damage limited to the hippocampal region. *Hippocampus*, 9, 495-499.
- Marr, D. (1971). Simple memory: a theory for archicortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 262, 23-81.
- Moll, M. & Miikkulainen, R. (1997). Convergence-zone episodic memory: Analysis and simulations. *Neural Networks*, 10, 1017-1036.
- Morton, J. & Bekerian, D. A. (1986). Three ways of looking at memory. In N. E. Sharkey (Ed.), *Advances in Cognitive Science 1* (pp. 43-71). Ellis Horwood Ltd.
- Morton, J., Hammersley, R. H., & Bekerian, D. A. (1985). Headed records: a model for memory and its failure. *Cognition*, 20, 1-23.
- Murre, J. M. (1996). Trace link: a model of amnesia and consolidation of memory. *Hippocampus*, 6, 675-684.
- O'Keefe, J. & Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford: Oxford University Press.

- Papez, J. W. (1937). A proposed mechanism of emotion. *Archives of Neurology and Psychiatry*, 38, 724-744.
- Proust, M. (Scott Moncrieff CK, transl.) (1922/1960). *Swann's Way*. London: Chatto & Windus.
- Roediger, H. L. & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 803-814.
- Rabin, M. D. & Cain, W. S. (1984). Odor recognition: familiarity, identifiability, and encoding consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(2), 316-325.
- Rubin, D. C., Groth, E., & Goldsmith, D. J. (1984). Olfactory cueing of autobiographical memory. *American Journal of Psychology*, 97, 493-507.
- Rugg, M. D. & Yonelinas, A. P. (2003). Human recognition memory: a cognitive neuroscience perspective. *Trends in Cognitive Sciences*, 7, 313-319.
- Rusted, J., Gaskell, M., Watts, S., & Sheppard, L. (2000). People with dementia use schemata to support episodic memory. *Dementia and Geriatric Cognitive Disorders*, 11, 350-356.
- Simons, J. S. & Spiers, H. J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience*, 4, 637-648.
- Spiers, H. J. (2002). *Temporal Lobe Contributions to Topographical and Episodic Memory*. Doctoral Thesis. UCL.
- Spiers, H. J., Burgess, N., Hartley, T., Vargha-Khadem, F., & O'Keefe, J. (2001a). Bilateral hippocampal pathology impairs topographical and episodic memory but not visual pattern matching. *Hippocampus*, 11, 715-725.
- Spiers, H. J., Burgess, N., Maguire, E. A., Baxendale, S. A., Hartley, T., Thompson, P., & O'Keefe, J. (2001b). Unilateral temporal lobectomy patients show lateralised topographical and episodic memory deficits in a virtual town. *Brain*, 124, 2476-2489.
- Spiers, H. J., Maguire, E. A., & Burgess, N. (2001c). Hippocampal amnesia. *Neurocase*, 7, 357-382.
- Squire, L. R. & Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science*, 253, 1380-1386.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory* (pp. 381-403). New York: Academic Press.
- Tulving, E. (1983). *Elements of episodic memory*. Oxford: Clarendon Press.
- Tulving, E. (2001). Episodic memory and common sense: how far apart? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 356, 1505-1515.
- Tulving, E. (2002). Episodic memory: from mind to brain. *Annual Review of Psychology*, 53, 1-25.
- Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., Van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science*, 277, 376-380.
- Wagenaar, W. A. (1986). My memory: A study of autobiographical memory over six years. *Cognitive Psychology*, 18, 225-252.

Wan, H., Aggleton, J. P., & Brown, M. W. (1999). Different contributions of the hippocampus and perirhinal cortex to recognition memory. *The Journal of Neuroscience*, *19*, 1142-1148.

Wheeler, M. A., Stuss, D. T., & Tulving, E. (1997). Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness. *Psychological Bulletin*, *121*, 331-354.

Wichawut, C., & Martin, E. (1971). Independence of A-B and A-C associations in retroaction. *Journal of Verbal Learning and Verbal Behavior*, *10*, 316-321.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441-517.

## Tables

Participant(s)	Question-type		
	Context-free	Context-dependent	
	Object	Person	Place
Jon	85%	60%	55%
control group average	86%	83%	84%
control group s.d.	6%	12%	9%

**Table 1** Hippocampal patient Jon's performance shows a comparable score for a purely familiarity-based recognition task (context-free 'Object'-condition) as control participants, as opposed to impaired performance in two context-dependent, episodic-memory-recognition tasks ('Person'- and 'Place'-condition). See King et al. (2004) for details.

Retrieval with another cue	Retrieval with one cue (more successful)	
	Proportion correct (f)	Proportion incorrect (1-f)
Proportion correct (p)	$a = p$	$b = 0$
Proportion incorrect (1-p)	$c = f-p$	$d = (1-f)$

**Table 2** Contingency table for a Fully Dependent Model, see text for explanations.

Retrieval with another cue	Retrieval with one cue (more successful)	
	Proportion correct (f)	Proportion incorrect (1-f)
Proportion correct (p)	$a = p * f$	$b = p * (1-f)$
Proportion incorrect (1-p)	$c = f * (1-p)$	$d = (1-p) * (1-f)$

**Table 3** Contingency table for the Independent Model, see text for explanations.

Comparison	Difference from Independent Model, P values ( $\chi^2$ statistic)	Difference from Dependent Model, P values ( $\chi^2$ statistic)
Object vs Person	0.09 (0.99)	0.06 (7.38)
Object vs First	0.99 (0.07)	0.02 (9.66)
Object vs Place	0.99 (0.01)	<0.001 (30.79)
Person vs Place	0.99 (0.09)	0.001 (20.34)
First vs Person	0.95 (0.34)	0.07 (7.18)
First vs Place	0.99 (0.09)	<0.001 (37.18)

**Table 4** Chi-square analysis for comparison of empirical data from Experiment 1 with both Dependent and Independent Models (see Tables 2 and 3).

Participant number	P value Object v Person	P value Object v Place	P value Person v Place
1	1.00	1.00	1.00
2	0.85	0.72	0.90
3	0.49	0.40	0.25
4	0.63	0.40	0.60
5	0.63	0.38	0.63
6	1.00	0.81	1.00
7	0.63	0.80	0.90
8	0.08	0.63	0.34
9	0.02	0.28	0.52
10	0.21	0.34	0.61
11	0.80	0.75	0.72
12	0.75	0.80	0.72

**Table 5** P-values (Fisher's Exact Test) for  $h_0$ =rejection of Independent Model per participant and event-element-pairing in Experiment 2. There is no sign of similarity in performance on different questions about the same event. Note that, for our performance levels (e.g. 0.85), a Fully Dependent Model With Guessing (see Table 6) would score  $p=0.15$  – i.e. we do not have the power to reject the Independent Model at  $p<0.05$ . However, the average p-values are clearly greater than 0.15, consistent with a Dependent Model.

Retrieval with another cue	Retrieval with one cue	
	Proportion correct ( $p = p' + g/2$ )	Proportion incorrect ( $1 - p$ )
Proportion correct ( $p = p' + g/2$ )	$a = p' + g/4$	$b = g/4$
Proportion incorrect ( $1 - p$ )	$c = g/4$	$d = 1 - p' - 3g/4 = g/4$

**Table 6** Contingency table for a Fully Dependent Model With Guessing in the case of equally good retrieval via either cue (as for Experiment 2) – proportion correct is  $p$  in both cases. The proportion of events in which both cues are correctly retrieved from memory is  $p'$ , the proportion of guessed answers is  $g$ . All responses are either due to correct retrieval of both aspects or due to random guessing, i.e.  $p' + g = 1$ .

Comparison	Independent Model P value ( $\chi^2$ , 1 d.o.f.)	Dependent Model With Guessing P value ( $\chi^2$ , 1 d.o.f.)
Object vs Person	0.92 (0.009)	<0.0001 (28.2)
Object vs Place	0.23 (1.461)	<0.0001 (39.2)
Person vs Place	0.099 (2.717)	<0.0001 (15.2)

**Table 7a** P-values and Chi Square statistics for both, the comparison of data with the expectation according to the Independent Model and the Dependent Model With Guessing over all participants (n=240). The Dependent Model can be clearly rejected whereas the Independent Model provides a reasonable fit.

Comparison	Mean squared difference (average over participants, sd in brackets) between data and..		P-value(t-test of mean squared diffs)
	..Independent Model	..Dependent Model With Guessing	
Object vs Person	0.0018 (0.0003)	0.0065 (0.0003)	<0.005
Object vs Place	0.0007 (5.6x10 <sup>-5</sup> )	0.0062 (0.0003)	<0.001
Person vs Place	0.0004 (6.6x10 <sup>-5</sup> )	0.0031 (0.0002)	<0.005

**Table 7b** Direct model-data comparison. For all three comparisons the mean squared difference  $[(a-a')^2+(b-b')^2+(c-c')^2+(d-d')^2]/4$  between model and data is significantly smaller for the *Independent Model* than the *Dependent Model With Guessing*.

	Question Type			
	Object	Odor	Person	Place
Average performance	100%	64%	80%	71%
(stdev)	(0%)	(17%)	(14%)	(22%)

**Table 8** Overall performance in the paired forced-choice test of Experiment 3 by question type, see text for details.

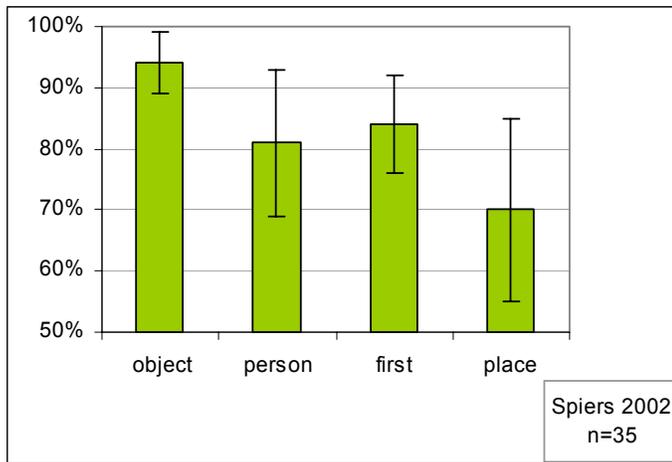
No. of combined elements remembered	Initially remembered	Eventually remembered
4 elements	7%	19%
3 elements	18%	29%
2 elements	41%	33%
<2 elements (i.e. forgotten)	34%	19%

**Table 9** Proportions of correctly remembered elements over all events and all participants in the Cued Recall task of Experiment 3.

3 elements remembered	Missing element				Total
	Object	Odor	Person	Place	
Initially	0%	16.4%	0%	1.6%	18%
Eventually	1.7%	23.2%	3.2%	0.9%	29%

**Table 10** Percentage of events of which all elements but one are reconstructed correctly, by omitted element. Odor is most frequently omitted.

## Figures



**Figure 1** Average performance over all participants per question type in Experiment 1. Error bars show 1 standard deviation. 'object' refers to a question solvable by familiarity-based recognition, 'person', 'place', and 'first' refer to questions requiring context-dependent memory.



**Figure 2** Snapshot of the encoding phase of Experiment 2. The participant follows a trail of green dots (a, the next dot to move over is colored red), encounters a distinct person in a distinct location and is presented with the image of a distinct object (b).



**Figure 3** An example of a context-dependent paired forced choice question from Experiment 2, showing a 'Place' question: "Which object did you see in this place?"