

Chapter 3

Part or parcel? Contextual binding of events in episodic memory

Iris Trinkler, John King, Hugo Spiers, and
Neil Burgess

Introduction

In our daily life we experience a vast number 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). Tulving (1983, p. 223) has suggested that the ‘prototypical unit of an episodic memory is an event’, and the different elements of an event are believed 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 spatiotemporal context are remembered together. In contrast, if this viewpoint were completely untrue, memory for different elements of an event might be remembered or forgotten *independently*. 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 the theoretical stand point of Tulving and 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 phenomenologically known to all healthy people who can

“travel back in time in their own minds” (Duzel *et al.* 1997, p. 5973). Tulving and his colleagues are not alone in their suggestion that events might be the units of episodic memory. Fisher and Chandler 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’ (Fisher and Chandler 1991, p. 722, emphasis added), 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 films 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 coloured objects, each in a specific location within the scene. They were then cued with the colour, 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 these data. First, 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 with its usefulness as a retrieval cue (in addition, retrieval of serial position decreased with serial position, while retrieval of other information from serial position followed a U-shaped curve, being best near to the start or end of a sequence). In contrast, other elements were used symmetrically: retrieval of element A by element B was as good as retrieval of element B by element A. This symmetrical use of cues was 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 pairwise association contributes independently to the probability of success. Jones suggested that memories of the visual characteristics of his events (object, colour, 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 4 years by probing himself with different elements of each event (who, what, when, and where). Consistent with

Jones's (1976) study, he found that temporal information (when) was a very poor cue even though it could be retrieved reasonably well. However, unlike 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 *et al.* 1985; Morton and Bekerian 1986). 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 interpreted the differences between his study and that of Jones in terms of differences in cue specificity. In Jones's study, within each list of nine events, each cue was specific to only one event, whereas Wagenaar's cues varied in specificity, with 'what' being 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's 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's stimuli in being truly multimodal. 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 (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 Fig. 3.2 below). 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), from the ongoing 'context' of the event, including the surrounding spatial environment, time, and the person giving the object (Burgess *et al.* 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 with a similar-looking novel foil. In one variation of this paradigm, we added odour 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 subtypes, 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 concerns the specific role in memory played by the hippocampus as often as it 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 (reviewed by Burgess *et al.* 2002).

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 (Kinsbourne and Wood 1975; O'Keefe and Nadel 1978; Squire and Zola-Morgan 1991; Vargha-Khadem *et al.* 1997; Eichenbaum and Cohen 2001; reviewed by Spiers *et al.* 2001c, Burgess *et al.* 2002), although 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 acontextual forms of memory such as memory for factual knowledge ('semantic memory'). Declarative memory theory (e.g. Squire and Zola-Morgan 1991) sees the medial temporal lobe (including the amygdala 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 (Mandler 1980, 1991, Jacoby *et al.* 1993, reviewed by Yonelinas 2002) 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') and 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 mammillary bodies, anterior thalamus, and hippocampus supporting episodic recollection, while a distinct parallel system, including the medial thalamus and perirhinal cortex, supports familiarity-based recognition (Delay and Brion 1969; Gaffan and Parker 1996; Vargha-Khadem *et al.* 1997; Aggleton and Brown 1999; Wan *et al.* 1999; Bogacz *et al.* 2001; Tulving 2001; Holdstock *et al.* 2002; Yonelinas 2002). In a recent review, Rugg and 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 (Spiers *et al.* 2001a, b; King *et al.* 2004). First, we discuss our experiments with Jon, a young man with focal bilateral hippocampal pathology (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 passed a UK General Certificate of Secondary Education (GCSE) examination 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 his context-dependent recognition memory was impaired. However, there were two potential problems with this study. First, 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

Table 3.1 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 conditions)

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

See King *et al.* (2004) for details.

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 (Table 3.1).

The pattern of performance shown by patient Jon conflicts with the declarative memory theory (Squire and 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 and Squire 1999; reviewed by Spiers *et al.* 2001c). In 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 the perirhinal cortex supports familiarity-based recognition (Gaffan and Parker 1996; Aggleton and Brown 1999; Baxter and 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 both these patients and 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 and Nadel's (1978, Chapter 14, p. 380 ff.) extension of the cognitive map theory to humans has more similarities with 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 (O'Keefe and Nadel, 1978, p. 384). Eichenbaum and 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 pairwise 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' (Damasio 1989; Alvarez and Squire 1994; Murre 1996; Moll and Miikkulainen 1997), linking information from different sensory modalities that are represented in disparate cortical areas. In Experiment 3, we added an olfactory component to our (visual) VR events in order to compare unimodal 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 (Chu and Downes 2000). Such a role might reflect the direct connections between primary olfactory regions and the hippocampus (Dade *et al.* 2002). However, scientific evidence rather undermined this notion (Bolger and Titchener 1907; Davis 1975; Rubin *et al.* 1984; Herz 1998). The only evidence for privileged olfactory cueing of memory comes from Chu and Downes (2002) who found that solely odour cues enhance autobiographical memory retrieval in a second retrieval attempt (following a first memory search cued by a label) compared with other cues. Alternatively, Rubin *et al.* (1984) argue that long-term retention in olfactory memory is due to odour 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 completeness, we should also mention that many areas outside 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 (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 (reviewed by Wheeler *et al.* 1997; Simons and Spiers 2003). 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 *et al.* 2001; King *et al.* 2005). 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 and Mummery 1999; Maguire *et al.* 2000) was interpreted as reflecting the interference resulting from such overlapping contextual cues. This interpretation is consistent with neuropsychological studies (Incisa della Rocchetta and 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.* 2005).

Experiment 1

In the first version of the VCM experiment, participants experienced a series of 16 events that took place in two different places within a VR environment, and with two different 'people' (VR 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 four times addressing memory for different aspects. The whole sequence was then repeated with 16 new objects in two new places with two new characters. In total, participants answered 128 memory questions on the VR events experienced. For further details see Spiers 2002.

Methods

Participants

Thirty-five participants (nine female, 26 male), with an age range of 18–33 years (mean 25 years), took part in the study. Their mean IQ was 105 (inferred from a mean score of 9.5 (SD 1.7) on Raven's Progressive Matrices, Set 1).

Encoding task

Participants followed a marked route through a VR 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 4×16 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.

- ◆ 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?'

The foil for the Object question was a similar looking version of the original, while foil objects in the other conditions were from other events in the encoding task and thus were equally familiar.

Practice trial

Participants were given a practice trial during which they followed each of the two routes and 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 simply to pay attention to the various elements of each event.

Results

There was a significant effect of question type on performance (Fig. 3.1). 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 the temporal cue First than via the Place cue.

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 3.2, and under a Fully Independent Model (where retrieving one element of an event is independent of retrieving another element of that event) in Table 3.3. 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

Figure 3.1 Average performance over all participants per question type in Experiment 1. Error bars show one standard deviation. Object refers to a question solvable by familiarity-based recognition; Person, Place, and First refer to questions requiring context-dependent memory.

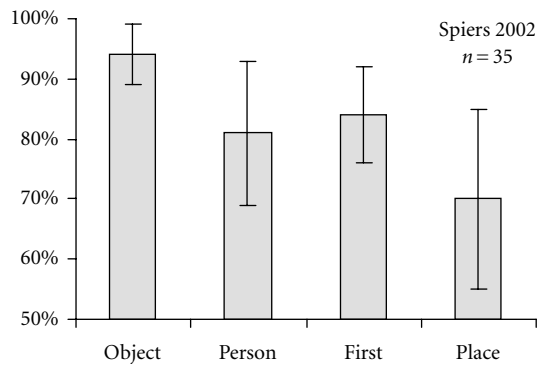


Table 3.2 Contingency table for a Fully Dependent Model

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

See text for explanation.

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 ($c = f - p$). In the Fully 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 Table 3.3).

Observed responses and predicted values from the Fully Dependent and Fully Independent Models were subjected to a χ^2 analysis. Table 3.4 shows the corresponding statistics and P -values. 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.

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

Table 3.3 Contingency table for the Independent Model

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

See text for explanation.

Table 3.4 χ^2 analysis for comparison of empirical data from Experiment 1 with both the Dependent and the Independent Models (see Tables 3.2 and 3.3)

| Comparison | P -value (χ^2 statistic) | |
|------------------|-----------------------------------|---------------------------------|
| | Difference from Independent Model | Difference from Dependent Model |
| 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) |

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 (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. Therefore random answers could 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 of 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 of retrieving 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

Twelve male participants, who were age and IQ matched for comparisons of performance with patient Jon (see above), took part in this experiment. Their

age range was 21–28 years (mean age 23.4) and their mean IQ was 114 (inferred from a mean score of 10.43 (SD 1.22) on Raven's Advanced Matrices, Set I).

Encoding task

A VR 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 800×600 pixels and a vertical refresh rate of 60 Hz. To manoeuvre within the town, participants used the cursor keys of the keyboard and followed a trail of green icons (Fig. 3.2(a)). In distinct places along the route, participants encountered virtual characters who presented them with an image of an object (display size 7×7 cm) (Fig. 3.2(b)). Subsequently, a new trail of icons would appear for the

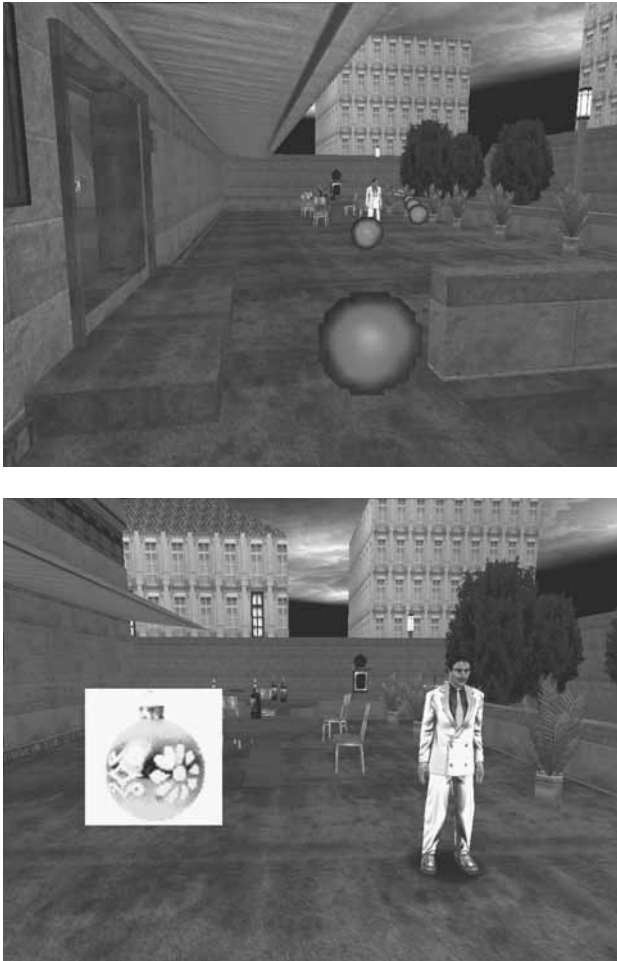


Figure 3.2 Snapshot of the encoding phase of Experiment 2. (a) The participant follows a trail of green dots (the next dot to move over is coloured red), and (b) encounters a distinct person in a distinct location and is presented with the image of a distinct object.

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 min on average.

Recognition task

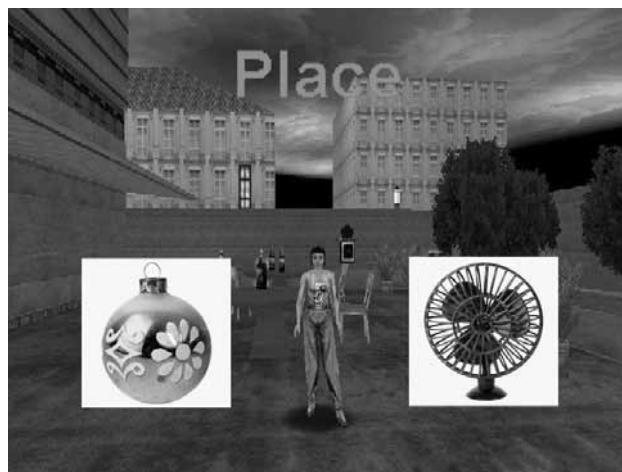
Immediately after the encoding task, participants were given paired forced-choice recognition tests on all aspects of all events (3×20 tests): they were presented with two objects, on the left and right of the screen, a virtual character in the foreground, and a snapshot of one of the locations in the background. A word appearing on the top of the screen indicated what type of event information was being probed (Fig. 3.3).

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

Before testing, participants were given a trial run of both the encoding task (consisting of three events presented in an alternative VR town) and the recognition memory test.

Figure 3.3 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?'



Results

The average performance over all participants is shown in Table 3.1. 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 analysed 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 Fisher's exact test was far from being rejected (Table 3.5).

We also explicitly created the contingency tables expected under an Independent Model (as explained in Experiment 1; see Table 3.3) on the 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 (Table 3.6).

Table 3.5 *P*-values (Fisher's exact test) for h_0 = rejection of Independent Model per participant and event-element-pairing in Experiment 2

| Participant no. | <i>P</i> -value | | |
|-----------------|------------------|-----------------|-----------------|
| | Object vs Person | Object vs Place | Person vs 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 |

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 3.6) 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 > 0.15 , consistent with a Dependent Model.

Table 3.6 Contingency table for a Fully Dependent Model With Guessing in the case of equally good retrieval via either cue (as for Experiment 2)

| 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 ($1 - p$) | $c = g/4$ | $d = 1 - p' - 3g/4 = g/4$ |

The proportion correct is p in both cases. The proportion of events in which both cues are correctly retrieved from memory is p' and the proportion of guessed answers is g . All responses are due to either correct retrieval of both aspects or random guessing, i.e. $p' + g = 1$.

Table 3.7 P -values and χ^2 statistics for comparison of data with the expectation according to both the Independent Model and the Dependent Model With Guessing over all participants ($n = 240$)

| Comparison | P-value (χ^2 , 1 degree of freedom) | |
|------------------|---|-------------------------------|
| | Independent Model | Dependent Model With Guessing |
| 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) |

The Dependent Model can clearly be rejected whereas the Independent Model provides a reasonable fit.

Table 3.8 Direct comparison of model and data

| Comparison | Mean squared difference ^a between data and | | P-value (t-test of mean squared differences) |
|------------------|---|-------------------------------|--|
| | Independent Model | Dependent Model With Guessing | |
| Object vs Person | 0.0018 (0.0003) | 0.0065 (0.0003) | < 0.005 |
| Object vs Place | 0.0007 (5.6×10^{-5}) | 0.0062 (0.0003) | < 0.001 |
| Person vs Place | 0.0004 (6.6×10^{-5}) | 0.0031 (0.0002) | < 0.005 |

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 for the Independent Model is significantly smaller than that for the Dependent Model With Guessing.

^a Average over participants; SD in parentheses

The χ^2 test applied to a contingency table over all participants corroborates data from the analysis of single participants (Table 3.7), fitting the Independent Model and rejecting the Dependent Model With Guessing.

Furthermore, we sought to compare the two models directly, evaluating the difference between both models and the data (Table 3.8). For all three pairs of questions compared, the sum of squared differences between model and data was significantly smaller for the Independent Model than for the Dependent Model With Guessing.

In summary, as in Experiment 1, we found no evidence in favour 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 pairwise 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 favour 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

Twelve participants (five females and seven males) with an average age of 28 years (range 22–39 years), took part in the experiment. Their mean IQ was 108 (inferred from a mean score of 9.75 (SD 2.0) on Raven's Advanced Matrices,

Set 1). Only people who rated their sense of smell 5 or above were included (self-rating on a scale from 1 to 7).

Practice trial

Before testing, participants were given the same trial run as in Experiment 2 with the addition of an odour 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 describe them verbally. 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 10 odour stimuli.

Encoding task

The virtual town, computer, and manner of navigating and responding were identical to Experiment 2. At 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 1 cm from the participant's nose) for the duration of one sniff. Participants then continued their journey through the virtual town along a new trail of green icons. They were told that they would be tested on these events subsequently and instructed to try to remember the person, object, place, and odour of each event.

Each participant experienced a unique composition of 10 events, i.e. assignment of objects, people, places, and odours was randomly varied between participants. There were five 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 about 15 min on average.

Odour stimuli

Odours were presented in medicine bottles labelled 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 were 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 event locations; in the Person condition, one of the 10 characters appeared between the two objects in front of a plain brown background; in the Odour condition, the participants were presented with one of the 10 odours 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 odour 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 online by the experimenter. We recorded which card-odour 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 3.9 shows the results of the recognition memory test in Experiment 3. Performance on the Object question was significantly better than performance on any other question type (repeated measures analysis of variance: overall

Table 3.9 Overall performance in the paired forced-choice test of Experiment 3 by question type

| | Question type | | | |
|-------------------------|---------------|-------|--------|-------|
| | Object | Odour | Person | Place |
| Average performance (%) | 100 | 64 | 80 | 71 |
| SD (%) | 0 | 17 | 14 | 22 |

See text for details.

$F = 15.7$, $df = 9$, $P < 0.001$; simple contrasts between Object and any other condition: F values of 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 Odour questions (paired t -tests, one-tailed, $P < 0.01$ for Person versus Odour, $P < 0.05$ for Person versus Place), but no significant difference between performance on Place and Odour questions.

Regarding the Odour cue, we assessed a possible relation between recognition memory score and odour identification ability: Odour 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 odour-cued recognition ($R^2 = 0.0048$); however, there was a high correlation between an odour's identification score and how well it elicited odour-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 per cent of all events (total = 10 events \times 12 participants) were at least partially correctly reconstructed as opposed to 34 per cent of events for which no two elements were correctly matched. For 'eventually remembered' items, these percentages amounted to 81 per cent at least partially correctly reconstructed events versus 19 per cent entirely forgotten (Table 3.10).

The percentage of events for which all four elements were remembered together correctly does not exceed 19 per cent. Thus the full retrieval of a complete event is rare compared with the retrieval of the object via a single cue in the recognition test (which was 64 per cent for the worst case, Odour). Table 3.11 shows which elements were not recombined with the others correctly when the other three elements were correctly matched, revealing that the association with odour was the weakest overall; it could not be reattributed to the other elements

Table 3.10 Proportions of correctly remembered elements over all events and all participants in the cued-recall task of Experiment 3

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

Table 3.11 Percentage of events in which all elements but one are reconstructed correctly, by omitted element

| Three elements remembered | Missing element | | | | |
|---------------------------|-----------------|-------|--------|-------|-------|
| | Object | Odour | Person | Place | Total |
| Initially | 0% | 16.4% | 0% | 1.6% | 18% |
| Eventually | 1.7% | 23.2% | 3.2% | 0.9% | 29% |

Odour is most frequently omitted.

of an event in 91 per cent 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 three-elements category, or participants may have taken the element combinations apart again and recombined them differently.

Events were retrieved via the following cues first: Person, 38 per cent; Object, 18 per cent; Odour, 7 per cent; Place, 5 per cent (note that these percentages do not sum to the total number of 66 per cent 'initially remembered' because some events were retrieved by two cues simultaneously). Participants thus show an overall cue preference in favour of the Person cue over the Place and Odour 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 at 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-coloured 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 either at an initial retrieval attempt or after the final retrieval attempt. However, the percentage of events for which more than two elements are correctly associated together increases from 25 per cent initially remembered to 48 per cent eventually remembered. Thus participants are 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) pairwise associations with either of the elements of the first pair.

In contrast with Experiment 2, which shared the VR setting of Experiment 3 (but not the odour element), performance in the Object condition was much higher than in the other (context-dependent) conditions of the recognition test. Possible explanations are as follows: first, the total number of events was only 10 rather than 20; second, the number of repeated exposures to objects in the recognition test was increased because of an additional Odour condition; thirdly, presentation of the odour distracted participants from the other elements of context. These last two factors may also have contributed to the worse performance in the context-dependent question types in this experiment. However, we note that we previously found that retrieval of different elements of an event was independent of whether difficulty was matched (Experiment 2) or not (Experiment 1).

Retrieval independence is complemented by retrieval asymmetry: The Person cue featured predominantly as the element by which other elements of an event were retrieved. In contrast, Odour was most noticeably forgotten and hardly served as a primary cue for the reconstruction of the event. Similarly, performance of the Odour cue in the forced-choice recognition test was worst. The success of a particular Odour cue stimulus was correlated with how nameable it was considered to be overall, thus indicating odour 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, which could in turn be due to a 'tip-of-the-nose-state' (Lawless and Engen 1977), but rather the relative nameability of the stimuli in general.

General discussion

We used a VR 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. In contrast, a model based on independent pairwise 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 the individual elements into the events that they had experienced (Experiment 3). We found that at the initial retrieval attempt only 7 per cent of all events were retrieved fully, with 59 per cent remembered partially and 33 per cent not remembered at all. Importantly, this performance increases to 19 per cent entirely remembered in the end, as opposed to 62 per cent partially remembered and 19 per cent 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 multimodal context in such a realistic way that a mechanism of autoeotic awareness (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 and 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 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 pairwise associations. In addition, he reports many events that were only partially remembered, and a failure to retrieve around 20 per cent of events (consistent with the final result in the cued recall of our Experiment 3).

Our results also contrast with Jones's (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's paradigm, using nine lists of each of nine objects in nine different colours 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 participants would not have had time to learn it as they performed 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 and 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 as 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 *et al.* 1985; Morton and Bekerian 1986) 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's (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's experiments perhaps had more similar, and reduced, semantic complexity than the elements of context in our experiment. In Jones's experiment, one element ('location') was one of nine grid positions on a backdrop and was thus similar to the other element ('colour') in its (lack of) semantic complexity. Asch and Ebenholtz (1962) similarly demonstrated approximate symmetry in the recall of two-component visual patterns but argued that asymmetry in other circumstances could be due to differential availability of the components, perhaps differential levels of semantic association.

Clayton *et al.* (2001), in their investigation of episodic-like memory in scrub jays, suggested that 'where' is the predominant element binding an episode together compared with '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. However, we hypothesize 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 for scrub jays caching food, for example, the most successful memory cues might well be places (triggering additional episodic information), for human participants wandering around in (VR) 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. The original paradigm used by Clayton and colleagues (Clayton and Dickinson 1998; Clayton *et al.* 2001) did not test memory for 'who' was involved in a caching event (but see Emery and Clayton 2001). An alternative explanation would be that cue preference is dependent on the distinctiveness of cues of the same category across events. For example, if 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 (e.g. imagine '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), and as shown in earlier experimental investigations of human memory (Jenkins and 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 and 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 odours, the success of olfactory cues in retrieving episodes was relatively small compared with 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 labelled some minutes ago (Engen and Ross 1973, Cain and Potts 1996), with participants sometimes becoming trapped in a 'tip-of-the-nose-state' (Lawless and Engen 1977). The binding of label and olfactory percept is volatile. Therefore, on the one hand, semantic integration of an odour enhances its success as a retrieval cue, as reported above and shown previously (Lawless and Engen 1977; Rabin and Cain 1984; Lyman and McDaniel 1990). 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 with other modalities (Lawless and 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 pairwise 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 (Kinsbourne and Wood 1975; Mishkin *et al.* 1997; O'Keefe and Nadel 1978; Squire and Zola-Morgan 1991; Eichenbaum and Cohen 2001). Thus our results 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 favourably via the Person cue as opposed to the Place, Object, or Odour 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.

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 with 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.

References

- Aggleton, J.P. and Brown, M.W. (1999). Episodic memory, amnesia, and the hippocampal–anterior thalamic axis. *Behavioral and Brain Sciences*, **22**, 425–490.

- Alvarez, P. and 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. and Ebenholtz, S.M. (1962). The principle of associative symmetry. *Proceedings of the American Philosophical Society*, **106**, 135–163.
- Baxter, M.G. and 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., and Giraud-Carrier, C. (2001). Model of familiarity discrimination in the perirhinal cortex. *Journal of Computational Neuroscience*, **10**, 5–23.
- Bolger, E.M. and Titchener, E.B. (1907). Some experiments on the association power of smells. *American Journal of Psychology*, **18**, 326–327.
- Brewer, W.F. and 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. and 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., and O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, **35**, 625–641.
- Burgess, P.W. and Shallice, T. (1996). Confabulation and the control of recollection. *Memory*, **4**, 359–411.
- Cain, W.S. and Potts, B.C. (1996). Switch and bait: probing the discriminative basis of odor identification via recognition memory. *Chemical Senses*, **21**, 35–44.
- Chu, S. and Downes, J.J. (2000). Odor-evoked autobiographical memories: psychological investigations of Proustian phenomena. *Chemical Senses*, **25**, 111–116.
- Chu, S. and Downes, J.J. (2002). Proust nose best: odors are better cues of autobiographical memory. *Memory and Cognition*, **30**, 511–518.
- Clayton, N.S. and 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., and 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., and 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. and Brion, S. (1969). *Le syndrome de Korsakoff*. Paris: Masson.

- Duzel, E., Yonelinas, A.P., Mangun, G.R., Heinze, H.J., and 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. and Cohen, N.J. (1988). Representation in the hippocampus. What do hippocampal neurons code? *Trends in Neurosciences*, **11**, 244–248.
- Eichenbaum, H. and Cohen, N.J. (2001). *From Conditioning to Conscious Recollection: Memory Systems of the Brain*. Oxford: Oxford University Press.
- Emery, N.J. and Clayton, N.S. (2001). Effects of experience and social context on prospective caching strategies by scrub jays. *Nature*, **414**, 443–446.
- Engen, T. and Ross, B.M. (1973). Long-term memory of odors with and without verbal description. *Journal of Environmental Psychology*, **100**, 221–227.
- Fisher, R.P. and 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. and 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., and Vargha-Khadem, F. (2000). Developmental amnesia associated with early hypoxic-ischaemic injury. *Brain*, **123**, 499–507.
- Gaffan, D. and Parker, A. (1996). Interaction of perirhinal cortex with the fornix-fimbria: memory for objects and ‘object-in-place’ memory. *Journal of Neuroscience*, **16**, 5864–5869.
- Hayman, C.A. and 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., *et al.* (2002). Under what conditions is recognition spared relative to recall after selective hippocampal damage in humans? *Hippocampus*, **12**, 341–351.
- Incisa della Rocchetta, A. and Milner, B. (1993). Strategic search and retrieval inhibition: the role of the frontal lobes. *Neuropsychologia*, **31**, 503–524.
- Jacoby, L.L., Toth, J.P., and Yonelinas, A.P. (1993). Separating conscious and unconscious influences of memory: measuring recollection. *Journal of Experimental Psychology: General*, **122**, 139–154.
- Jenkins, J.J. and 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., and 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., and Burgess, N. (2005) Anterior prefrontal involvement in episodic retrieval reflects contextual interference. *NeuroImage*, **28**, 256–267.

- Kinsbourne, M. and Wood, F. (1975). In *Short-Term Memory* (ed D. Deutsch and J.A. Deutsch). New York: Academic Press, pp 257–291.
- Lawless, H.T. and 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. and 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. and 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., and 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 *Relating Theory and Data: Essays on Human Memory in Honor of Bennet B. Murdock* (ed E. William, E. Hockley, and E. S. Lewandowsky). Hillsdale, NJ: Erlbaum, pp. 207–225.
- Manns, J.R. and 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.
- Mishkin *et al.* (1997).
- Moll, M. and Miikkulainen, R. (1997). Convergence-zone episodic memory: analysis and simulations. *Neural Networks*, **10**, 1017–1036.
- Morton, J. and Bekerian, D.A. (1986). Three ways of looking at memory. In *Advances in Cognitive Science 1* (ed N.E. Sharkey). Ellis Horwood, pp. 43–71.
- Morton, J., Hammersley, R.H., and 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. and 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. (1922). *Swann's Way* (trans C.K. Scott Moncrieff). London: Chatto and Windus, 1960.
- Rabin, M.D. and Cain, W.S. (1984). Odor recognition: familiarity, identifiability, and encoding consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **10**, 316–325.
- Roediger, H.L. and McDermott, K.B. (1995). Creating false memories: remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 803–814.
- Rubin, D.C., Groth, E., and Goldsmith, D.J. (1984). Olfactory cueing of autobiographical memory. *American Journal of Psychology*, **97**, 493–507.

- Rugg, M.D. and Yonelinas, A.P. (2003). Human recognition memory: a cognitive neuroscience perspective. *Trends in Cognitive Sciences*, **7**, 313–319.
- Rusted, J., Gaskell, M., Watts, S., and Sheppard, L. (2000). People with dementia use schemata to support episodic memory. *Dementia and Geriatric Cognitive Disorders*, **11**, 350–356.
- Simons, J.S. and Spiers, H.J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience*, **4**, 637–648.
- Spiers, H. (2002). Temporal lobe contributions to topographical and episodic memory. Doctoral Thesis. University of California at Los Angeles.
- Spiers, H.J., Burgess, N., Hartley, T., Vargha-Khadem, F., and 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., *et al.* (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., and Burgess, N. (2001c). Hippocampal amnesia. *Neurocase*, **7**, 357–382.
- Squire, L.R. and Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science*, **253**, 1380–1386.
- Tulving, E. (1972). Episodic and semantic memory. In *Organization of Memory* (ed E. Tulving and W. Donaldson). New York: Academic Press, pp. 381–403.
- 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.
- Tulving, E. and Markowitsch, H.J. (1998). Encoding of information into the episodic system depends critically on the semantic system, p.200.
- Vargha-Khadem, F., Gadian, D.G., Watkins, K.E., Connelly, A., Van Paesschen, W., and 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. and Brown, M.W. (1999). Different contributions of the hippocampus and perirhinal cortex to recognition memory. *Journal of Neuroscience*, **19**, 1142–1148.
- Wheeler, M.A., Stuss, D.T., and Tulving, E. (1997). Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness. *Psychological Bulletin*, **121**, 331–354.
- Wichawut, C. and 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.