Copy No: _________

Human Memory Models for Operator Simulation

Simon Dennis
University of Adelaide

Michael S Humphreys
University of Queensland

Scott W Boland
University of Queensland

Steven Savvas
University of Adelaide

Shayne D Loft
University of Queensland

Defence Research and Development Canada - Toronto

Contract Report

Dr. Simon Dennis, Department of Psychology, University of Adelaide.

Email: Simon.Dennis@adelaide.edu.au.

CR 2005-185

CSA, Peter Kwantes, (416) 635-2028

2005-08-29

Authors

on behalf of the authors.

S. Dennis, M Humphreys, S. Boland, S Savvas, S. Loft

Approved by

AH/SMART

K.C. Hendy

A/Head, SMART

Approved for release by

Kathy M. Sutton

Chair, Document Review Panel

Abstract

In this review of the human memory literature, we focus on those areas critical to human operator modeling and on the computational techniques from the memory modeling and machine learning literatures that are relevant to simulating human behavior in this area. We outline the areas of short-term memory, semantic memory, episodic memory, prospective memory and categorization. In addition, we review models that span these areas and memory phenomena which have yet to attract modelling efforts but which are likely to be important in operator modelling. Finally, we outline our Recommendations as to which areas will need specific attention in order to build robust models of human operators.

Résumé

Dans cette revue de la littérature sur la mémoire humaine, nous mettons l’accent sur les secteurs essentiels pour la modélisation de l’opérateur humain et sur les techniques de calcul tirées de la littérature sur la modélisation de la mémoire et l’apprentissage machine qui permettent de simuler le comportement humain dans ce domaine. Nous abordons les domaines de la mémoire à court terme, la mémoire sémantique, la mémoire épisodique, la mémoire prospective et la catégorisation. Nous examinons également les modèles qui couvrent les domaines et les phénomènes de mémoire qui n’ont pas encore fait l’objet de modélisations, mais qui seront probablement importants dans la modélisation de l’opérateur. Nous présentons enfin nos recommandations sur les domaines auxquels il faudra prêter une attention particulière pour construire de solides modèles d’opérateur humain.

This page intentionally left blank.

Executive summary

The human memory literature is one of the most extensive in cognitive psychology. Our starting point is the review of the human modeling literature conducted by Pew and Mavor (1998). This report was the work of the Panel on Modeling Human Behavior and Command Decision Making: Representations for Military Simulations. The panel was established by the National Research Council (NRC) in 1996 to review the state of the art in human behavior representation as applied to military simulations and the report contains a chapter devoted to memory and learning. Pew and Mavor (1998) argue that memory can be broken down into three broad categories: (1) episodic, generic, and implicit memory; (2) short-term and working memory; and (3) long-term memory and retrieval. In each of these sections, we have highlighted the range of psychological models that have been proposed, outlined briefly the mechanics of the most successful models, discussed the associated data in order to assess psychological plausibility and then discussed the relevance of the model for operator simulation. We suspect that the area of human operator modelling is about to undergo a transformation. Rather than the large finely crafted models that predominate today, we will see a generation of far simpler models that apply machine-learning techniques to large data sets of operator actions. To increase the fidelity of these models, it will continue to be important to consider key memory phenomena and models in the areas of short-term memory, semantic memory, episodic memory, prospective memory and categorization. Furthermore, we believe that human operator modeling provides a useful context to guide future empirical work and the next generation of memory models.

Dennis, S., Humphreys, M., Boland, S., Savvas, S., Loft, S. 2005. Human Memory Models for Operator Simluation. CR 2005-185 DRDC Toronto.

Sommaire

La littérature sur la mémoire humaine est une des plus complètes qui soit en psychologie cognitive. Notre point de départ est l'examen de la littérature sur la modélisation humaine effectué par Pew et Mavor (1998). Ce rapport a été fait par le Panel on Modeling Human Behavior and Command Decision Making: Representations for Military Simulations. Le groupe a été créé par le National Research Council (NRC) en 1996 dans le but d'examiner la représentation la plus récente du comportement humain appliquée aux simulations militaires, et le rapport contient un chapitre sur la mémoire et l'apprentissage. Selon Pew et Mavor (1998), la mémoire se divise en trois grandes catégories : (1) la mémoire épisodique, générique et implicite; (2) la mémoire à court terme et la mémoire de travail; et (3) la mémoire à long terme et la récupération. Dans chacune de ces sections, nous avons présenté la gamme des modèles psychologiques proposés, souligné brièvement les rouages des modèles les plus fructueux, présenté les données connexes en vue d'évaluer la vraisemblance psychologique puis étudié la pertinence du modèle pour la simulation de l'opérateur. Nous croyons que le domaine de la modélisation de l'opérateur humain est à la veille d'une transformation. Les modèles finement ciselés qui prédominent aujourd'hui céderont la place à une génération de modèles beaucoup plus simples qui appliqueront des techniques d'apprentissage machine à de grands ensembles de données sur les actes posés par les opérateurs. Il demeurera important, pour rendre ces modèles plus fidèles, d'étudier les phénomènes et les modèles clés de la mémoire à court terme, de la mémoire sémantique, de la mémoire épisodique, de la mémoire prospective et de la catégorisation. Nous croyons également que la modélisation de l'opérateur humain permettra d'orienter le travail empirique ultérieur et la prochaine génération de modèles de mémoire.

Dennis, S., Humphreys, M., Boland, S., Savvas, S., Loft, S. 2005. Human Memory Models for Operator Simluation. CR 2005-185 DRDC Toronto.

Table of contents

List of figures

This page intentionally left blank.

0. Introduction

The human memory literature is one of the most extensive in cognitive psychology. There have been a number of recent reviews of this literature (Nairne, 2002; Healy & McNamara, 1996; Bower, 2000; Raaijmakers & Shiffrin, 2002; Neath & Surprenant, 2003), however, any summary must necessarily omit a great deal. In this review, we focus on those areas critical to human operator modeling and on the computational techniques from the memory modeling and machine learning literatures that are relevant to simulating human behavior in this domain. Our starting point is the review of the human modeling literature conducted by Pew and Mavor (1998). This report was the work of the Panel on Modeling Human Behavior and Command Decision Making: Representations for Military Simulations. The panel was established by the National Research Council (NRC) in 1996 to review the state of the art in human behavior representation as applied to military simulations and the report contains a chapter devoted to memory and learning. Pew and Mavor (1998) argue that memory can be broken down into three broad categories: (1) episodic, generic, and implicit memory; (2) short-term and working memory; and (3) long-term memory and retrieval. While we concur that these categories cover the breadth of work that might be relevant and we have used this review to ensure that we have covered critical areas, we felt that the division used in Pew and Mavor (1998) was somewhat idiosyncratic. In particular, the first and third categories seem to overlap significantly. Instead, we have divided the area of memory into short-term memory, semantic memory, episodic memory and prospective memory. In addition, we have included a section on categorization, which is very closely related and is clearly relevant to the performance of human operators. In each of these sections, we have highlighted the range of psychological models that have been proposed, outlined briefly the mechanics of the most successful models, discussed the associated data in order to assess psychological plausibility and then discussed the relevance of the model for operator simulation. Finally, we discuss future directions - the areas in which we believe effort will need to be focused in order to build robust and reliable models of human performance.

1. Short-Term Memory

Memory for serial order is clearly critical in understanding and modelling human operators. Many tasks require serial order information to be retained and error can often be attributed to breakdown in short-term memory performance. The area of short term memory and, in particular, serial recall has engendered a great deal of debate and led to the creation of a number of simple, yet powerful computational models. Current accounts can be divided into chaining models, ordinal models and positional models. Before examining these models in detail however, we will overview the key data for which these models should account.

1.1 Effects/Paradigms of Serial Order Memory

The typical form of a serial learning experiment involves presenting a list of items one at a time, and then recalling that list immediately. This is called the immediate serial recall task. Two major variations of the immediate serial recall task are anticipation tasks and probing tasks. The anticipation task is similar to the typical form in that each item in the list is presented one at a time. However, recall is different as an item is shown first and the task is to anticipate the next item by recalling it before it too is shown. The next item is then shown, whether recall was correct or not. In probing tasks participants are given a position cue and asked to recall the item that appeared in the position.

Serial learning curves: Figure 1 illustrates the main features of serial learning items at the start of the list have the highest probability of recall (the primacy effect), this dips markedly through recall of the middle items, before increasing again at the end of the list (recency effect). Note that the recency effect is more pronounced in immediate serial tasks than in anticipation tasks shown here. The primacy effect is linked with covert/subvocal rehearsal and vanishes when the opportunity for rehearsal is removed (such as with articulatory suppression experiments).

Hebb effect: Unsurprisingly, learning takes place over many trials (see Figure 1: trial 1, 5 and 9). Items in the middle positions of a list also show comparatively greater improvement in recall then the start and the end of the list over many trials. Collectively these are known as the Hebb effect (Hebb, 1961).

Partial report effects: In experiments employing the probing paradigm results are still u-shaped, but show stronger recency than primacy (Murdock, 1968).

Delayed recall effects and proactive interference: Delayed recall effects are shown by inserting delay times/tasks between presentation of a list and its recall (the time between the end of list presentation and the beginning of recall is called the retention interval). Counting backwards is an example of a distracter task. Recall performance declines as retention interval increases. Conversely, if the retention interval is held constant across many trials, recall deteriorates due to the build-up of proactive interference from prior lists (Keppel & Underwood, 1962).

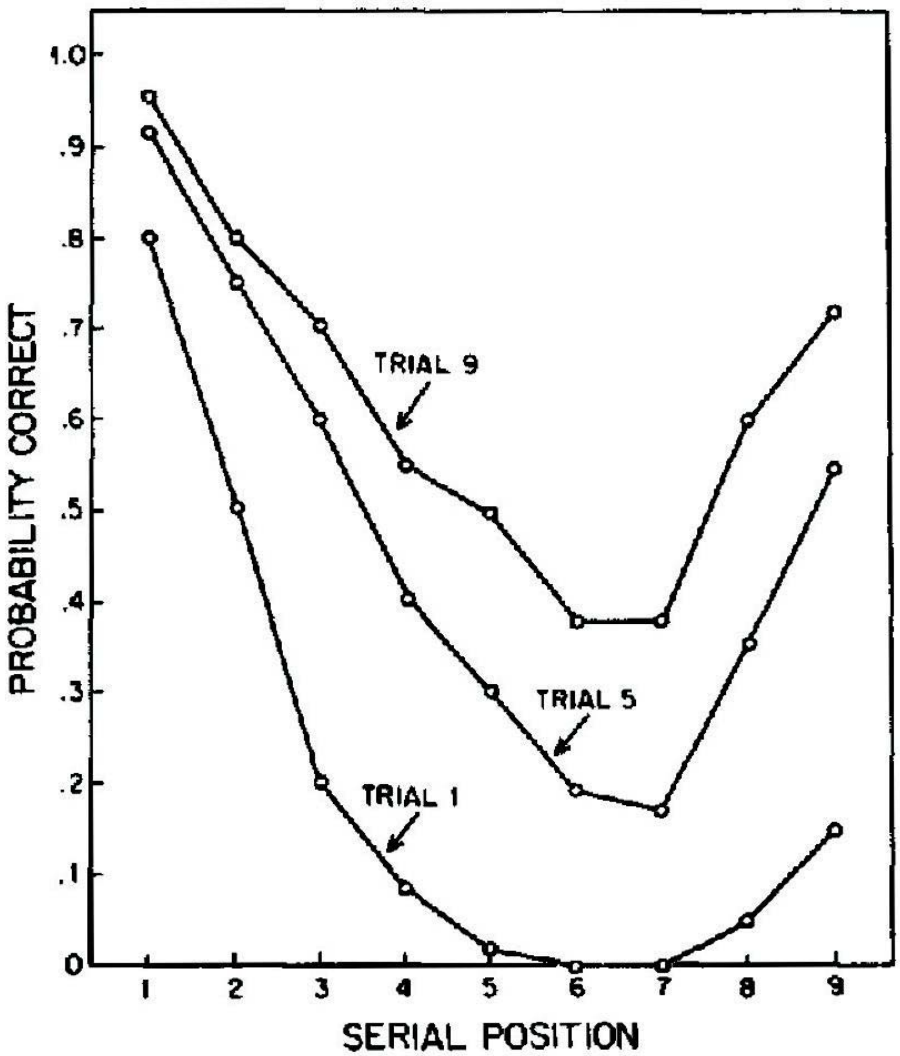

Figure 1. Anticipation task data for serial learning - from Lewandowsky and Murdock (1989).

Error patterns: Errors in recalling items in lists are not random, but rather adhere to certain patterns. Transpositions are common, where the order of consecutive items has been erroneously switched during recall. It is found that if an item is recalled in the wrong position, than it is more likely that it will be recalled in a position near its original position then in a position more distant. Inter-list intrusions occur when an item from another list intrudes in the recall of the current list, often in the same position (Estes, 1991).

Lag-recency effect: Another effect observed is that when an item is recalled correctly, the next item recalled tends to come from a nearby position (Kahana, 1996). This is referred to as the lag-recency effect. This effect has a characteristic shape (see Figure 2). Given that the current item was recalled correctly, it is most likely that the next item recalled will be the item that follows the current item (which has a lag of one). More remote items (which have larger lags) have a smaller chance of being recalled next. The figure also shows that the lag-recency effect is asymmetric recalling items in the forward direction is more likely than in the backward direction.

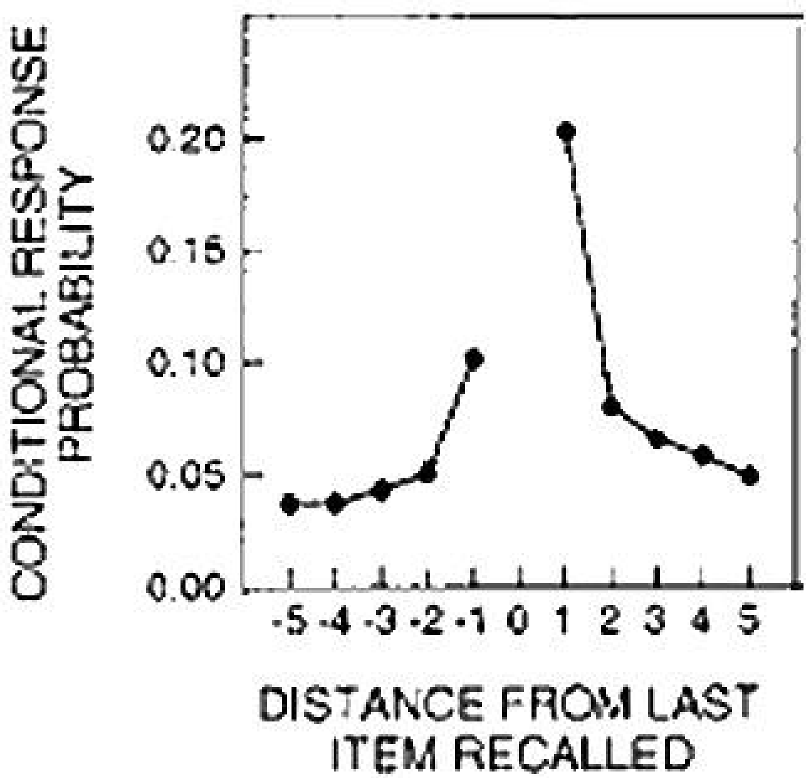

Figure 2. Free recall task data for lag-recency effect. Reproduced from Kahana (1996).

Phonological similarity effect: The phonological similarity effect is observed when performance for recalling items that sound similar is worse than recalling items that don't sound similar (Baddeley & Hitch, 1974).

Word-length effect: The word-length effect is an effect where lists of short words (in terms of duration to speak them) are easier to recall than lists of long words (Baddeley, Thomson, & Buchanan, 1975). The faster the articulation rate, the better the recall performance. However, when articulatory suppression tasks are employed, recall performance suffers and the word-length effect disappears (Baddeley & Wilson, 1984). Many models have been theorized to account for memory for serial order. Outlined next are the three main models types - chaining models, ordinal models, and positional models. These models all have strengths and weaknesses in accounting for the serial memory effects just described.

1.2 Chaining Models

Chaining models share the common feature that during learning, each item in a list is associated with the next item (like a link of chains). Therefore each element is the cue for recalling the next element in the list. The simplest chaining models associate only adjacent items, creating pair-wise associations. A series of these associations stores the serial order information. More complicated models allow for associations to form between non-adjacent items so that there may be multiple cues for the recall of the next item in a list.

Lewandowsky and Murdock's (1989) model is an extension of the Theory of Distributed Associative Memory (Murdock, 1983). Murdock's Theory of Distributed Associative Memory (TODAM) is a distributed memory model system where all information (whether that be item information or associations between items) is stored in a common pool. Associations are represented by the convolution operation and retrieval is achieved using the correlation operation. Lewandowsky and Murdock's model extends TODAM to account for serial recall.

Simple chaining models find it difficult to account for how items are retrieved if an item is recalled incorrectly as the chain has now been broken. Lewandowsky and Murdock's extension to TODAM circumvents this by assuming that an approximation of the preceding item is always available as a cue. That is, a previous item is used as a cue only if it is retrieved with clarity. Otherwise the approximation of the item that was retrieved from memory is used.

Chaining models have the property that retrieval of items in the middle of a list is dependent in some way on determining the items preceding it. This is also a property of next class of memory models for serial order called Ordinal models.

1.3 Ordinal Models

Ordinal models don't rely on the creation of associations between items (like previous chaining models do). Instead, order is stored in terms of the relative activation levels of each item in the list. Thus each node item in memory has an activation strength. The rank order of these strengths represents the list ordering. This activation pattern is called the primacy gradient.

With the Primacy model (Page & Norris, 1998), as each item is presented at study, an item node is activated to a certain degree to store the item in memory. This level of activation decreases linearly as more items are presented for learning, creating a primacy gradient. During recall, the item with the largest activation level is selected and then subsequently suppressed so it is not recalled again. This process repeats until all list items have been recalled. Activations are also subject to exponential time-based decay. This mechanism is needed to accommodate the common finding that errors in recall increase with list length.

Noise (which is modeled only in the recall process and not the learning process for simplicity sake) is introduced in the model to explain the common features in errors in serial learning such as transposition errors. The model accounts for the word-length effect (where short word lists are easier to recall then long word lists) by suggesting that subvocal rehearsal is simply a re-presentation of the list, returning items to their original strengths. Therefore, the strength of the primacy gradient is based on the last re-presentation of the list (i.e. the last time subvocal rehearsal occurred), and not based from the time the first item in the list was presented (at the start). Subvocal rehearsal continues in this manner until the list length is too large for subvocal rehearsal to occur in the time between item presentations. This is how the model accounts for the word-length effect as lists with shorter words have more opportunities to be rehearsed during subvocal rehearsal. This also explains the effects of list length on recall as shorter lists (lists with few words) have more opportunity for subvocal rehearsal than long lists and are therefore easier to learn.

Transpositions and repetition errors are handled without complication, but the Primacy model loses its elegance when accounting for the phonological similarity effect (where phonologically similar lists are recalled worse then lists that do not have similar sounding words). To deal with this result, the primacy model adopts a 2-stage mechanism, adding a confusion stage' to the model. This second stage incorporates the chance that phonologically similar items will be confused. This alteration models real world data well, but at the expense of model simplicity. The Primacy model also has problems dealing with positional intrusions (where items from previous leaned lists intrude in the same positions in the recall of a current list). Positional models (outlined next) account easily for these sorts of errors.

Like Chaining models, Ordinal models are dependent to a degree on determining the items at the start of a list before retrieving the middle items. Positional models are generally not restricted in this way, and can access middle items directly by reinstating the positional cue.

1.4 Positional models

In positional models, each item is associated with a unique positional cue. Bin approach models where items are put into position bins or slots have been replaced in the main by models that associate each item with a context. Models that use contexts include the Oscillator-based Associative Recall Model or OSCAR model (G. D. A. Brown, Preece, & Hulme, 2000), the connectionist model of the phonological loop (Hitch, Burgess, Towse, & Culpin, 1996), the Start-End model (Henson, 1998), the temporal context model (Howard & Kahana, 2002), and the neuropsychological motivated approach (Davelaar, Goshen-Gottstein, Ashkenazi, & Haarmann, 2005). The production system model based on Adaptive Control of Thought-Rational theory (Anderson & Matessa, 1997) is also a Positional model but it does not use contexts. It is more akin to a slot model with a hierarchical component.

With the Oscillator-based Associative Recall model (G. D. A. Brown et al., 2000) each list item is associated with a learning-context signal that is in a constant state of change. When a list is presented, an association is learned between an item and a particular learning context signal (whatever it happens to be at the time). The context signal is a combination of a series of oscillators that range from slow to fast. As the context signal changes during list presentation, each item in the list is associated with a unique learning context vector.

Recall of an individual item involves reinstating the learning context vector to the state it was in for that item. Here the model assumes that the learning context vector has intrinsic properties that allow it to be constructed by knowing simply the initial state (such as the first item in the list). Thus the entire list can be recalled by extrapolating from this initial state.

The model copes well with accounting for a host of serial order effects, especially item and order memory differences and list length effects. However, it is beyond the scope of the model to account for active rehearsal effects during list presentation.

The connectionist model of the phonological loop (Hitch et al., 1996) also invokes the use of a context signal timing in a similar fashion to the oscillator model just described. This context layer (composed of a set of nodes) is also associated with each representation of an item in memory. The difference here is that the context signal is modeled as a moving window' such that each context node is activated for only a short amount of time and that adjacent nodes are also partially activated. When recalling, the context representations are reconstructed in order with each context node used to refer back to an item in the item layer. After a context node is used to cue an item, it is given an inhibitory decay to avoid it being recalled again. Like OSCAR, this model also proposes that the context signal is generated by a set of temporal oscillators.

This model reproduces the classic u-shaped serial position curve. It explains error in recall as occurring as a consequence of either decay of the associations between an item and its context node, or noise' interfering with the activation of an item node. Errors due to association decay provide explanations for the word-length effect. However, the model has been criticized for its use of decay between context and item nodes to explain word length effects, as it does not take into account the contribution of rehearsal effects.

The Start-End Model (Henson, 1998) is similar to the connectionist model of the phonological loop, except that the Start-End model codes the position of list items relative to the start and end points of the list. The start point (or marker) is strongest at the start of the list, diminishing as it progresses to the end of the list. Conversely, the end marker is weakest at the start of the list, gaining maximum strength at the end of the list. The question arises as to how early items can be coded relative to the end marker when the end of the list hasn't even been presented yet.

The model provides the explanation that the strength of the end marker corresponds to the degree of expectation for the end of the list. The model proposes that when an item is presented for learning, a token is created in short term memory that gives an item a particular localized context. Each token can be seen as a code defining position relative to the start and end markers. Token ordering occurs during recall. To cue a particular list position, the model reinstates the code of that position and compares in parallel all the tokens against this code for matches. Items similar to the code compete for output. Once an item has been outputted, it is inhibited to prevent repetition errors.

The Start-End model can reproduce many serial order effects such as serial position curves, list length effects, and errors (transpositions, repetitions and omissions). It can also handle proactive interference and be extended to explain articulatory suppression tasks. As mentioned above, the greatest weakness of this model is that the end marker is used to code early list items even though the end of the list hasn't yet been seen. The explanation that the degree of expectation for the end of the list (i.e. the expected end of list length) is suitable only for controlled experiments where the list length is known or easily deduced. However, this cannot be assumed for many situations.

A production system theory of Serial Memory (Anderson & Matessa, 1997) does not make use of contexts but is still a positional model. This model is based on Adaptive Control of Thought-Rational (ACT-R) theory (Anderson, 1993). This model makes certain assumptions about serial lists - mainly that a list is organized in a sequence of groups and that each group is comprised of a series of items. As an example, consider a telephone number (made up of nine digits). This model assumes that the list is subdivided into groups (say three groups with three items each). At presentation of the telephone number, each digit is encoded in memory as an element with a value and also as having a position in the higher order structure (here the first digit would be item one in the first group, and the fifth digit would be item two in the second group). Thus each item is indexed by its position rather than items being linked by associations (such as in chaining models or context based positional models).

This approach models memory for serial order effects such as the word-length effect, the decline in recall with increasing list length, and errors such as transpositions, and most of the effects in free recall tasks (such as the primacy effect). However, the model consistently underestimates the significance of recency effects in free recall tasks. This model also does not attempt to model the effects of rehearsal during list presentation.

The Temporal Context model (Howard & Kahana, 2002) is a model based on context drift (another positional model that associates context with items). In this case the context (which is a set of binary elements) changes constantly in what is called contextual drift. Unlike other contextual drift proposals (Mensink & Raaijmakers, 1988), in TCM each new context is constructed by using the current item as a retreival cue and incorporating the context vector that is retrieved into the current context representation. The retrieval process for TCM involves re-invoking the context to the state it was when the item was studied. This is the cue for item access. This retrieved context state is then used for subsequent recalls (as it has more elements in common with other context states about this recalled state).

Each element in the context vector is associated with a particular strength to each element in the item vector (by way of a matrix of association strengths). It is assumed that this matrix is reset at the beginning of the learning of each new list. Thus a context node is used as a cue to retrieve an item node. Another matrix allows cueing in the opposite direction. That is, presentation of an item allows that item node to retrieve the prior state of context at the time that item was previously presented. TCM provides explanations of recency effects and the lag-recency effect (where if an item is recalled correctly, the next item recalled tens to come from a nearby position) across time scales.

Another model was developed in response to the other models not being able to account simultaneously for the recency effect found in free recall trials and long-term recency effects found when distracter tasks were inserted between each item presentation (the continuous distracter task). The Neuropsychological motivated approach (Davelaar et al., 2005) is a combination of a context-retrieval component of a short-term store that functions via activation levels. In effect, this mixed model is a combination of a positional model (an episodic contextual system) and an ordinal model (an activation-based short term buffer that has a limited capacity).

As with other context-based models, the context (vector) is associated with each item that is presented and changes during the presentation of lists. This also means that items nearby in position also have similar contexts. During learning, each item is stored in the short term buffer at a certain activation level and is also associated with a weight’. These weights (conceptualized as a weight based matrix) connect items in the buffer with context units and depict the strength between these two connections.

The first step in retrieval in a free recall task is the unloading of the active buffer (items currently in the short term buffer). If there is more than one then the weights’ are used to determine which item has the strongest connection to the current context. This then determines output order. Items that have been selected still compete for selection again. If they are selected again though, they don’t produce an output. The context vector then changes (based on a simple random walk algorithm) and items associated with this new context compete for selection. This process continues until all items have been retrieved.

This approach models serial position functions in both immediate free recall tasks and continuous-distracter tasks. It also handles lag-recency effects and the dissociations between immediate free recall task patterns and continuous-distracter task patterns. However, this model predicts long-term recency effects lower in magnitude then what is found empirically.

1.5 Operator Modelling

Typically, when modeling short-term memory for operator simulation, emphasis is placed on capacity restrictions. An assumption is made that performance will be compromised if too much information must be integrated at once. While this is a valuable component to include, as outlined above, the short-term memory literature is extensive. There are many effects that are likely to be playing a role in operator performance, such as proactive interference effects and similarity effects that typically are not incorporated in current simulations. Furthermore, there are many models that could potentially be employed. Recent work has tended to focus on the general class of positional models, but these models come in many forms and no one model has been demonstrated to be clearly superior over the entire set of short-term memory phenomena. Further comparative work, will be necessary before a clear choice will emerge.

In the next section, we shift the discussion from the short term to the much longer term as we consider the area of semantic memory.

2. Semantic Memory

The area of semantics and semantic memory is central to cognitive science and the construction of faithful models of human operators. In any situation in which an operator must deal with a stream of information the way in which they interpret and organize that information will to a large extent determine their performance.

While semantics can be used in a general way to refer to the construction of any kind of meaning, the study of semantics has tended to focus on language1. In this section, we will start by outlining the three main views of semantics that have arisen in the philosophical and linguistics literatures, namely truth conditional semantics (Frege, 1879), conceptual semantics (Jackendoff, 1992) and cognitive grammar (Lakoff, 1987; Langacker, 1987). Conceptual semantics and cognitive grammar make a key distinction between lexical semantics (the meanings of words) and compositional semantics (the meaning of sentences/utterances). This distinction is mirrored in the psychological literature. Early work on semantic memory focused on the meaning of words evolving from the hierarchical model (M. Collins & Quillian, 1969), through spreading activation models (A. Collins & Loftus, 1975) to feature based models (E. E. Smith, Shoben, & Rips, 1974), while models such as ACT (Anderson, 1983, 1993; Anderson & Lebiere, 1998) and Construction Integration theory (Kintsch, 1998) have emphasized the need for propositional representations to capture the meaning of sentences. In this section, we will focus on more recent techniques that allow both lexical semantics and compositional semantics to be extracted using large text corpora (Landauer & Dumais, 1997; Dennis, 2005).

2.1 Perspectives on Semantics

There are three main approaches to semantics that have arisen in the philosophical and linguistics literatures. The oldest and probably the most broadly understood perspective is truth conditional semantics, which was developed primarily in the philosophical literature. Truth Conditional Semantics focuses on the meaning of expressions and builds on the distinction between the referent of an expression and the sense of an expression (Frege, 1879). The reference is the thing that the expression refers to at this particular time in this particular place, whereas the sense of an expression refers to the meaning of the expression across circumstances. So, the expression “the prime minister of Australia” currently refers to John Howard. John Howard is the reference of the expression. Presumably in twenty years time, he will no longer be the prime minister of Australia, so the reference will have changed, but the expression will still mean the same thing, that is, the sense will be the same. According to truth conditional semantics then the meaning of an expression is the set of references it could refer to in all possible worlds. Defining semantics in this way makes meaning a relationship between expressions and the world.

By contrast, conceptual semantics (Jackendoff 1992) takes meaning to be a relationship between language symbols and mental concepts (see figure 3. The meanings of concepts are assumed to be defined by features, some of which are semantic primitives that are provided by our innate endowment. In addition, images of objects and our knowledge of how to perform actions play a role in defining concepts.

To capture the way in which conceptual structures are formed Jackendoff (1992) refers to different types of information including conceptual formation rules that map sets of concepts to simpler equivalent concepts, inference rules that map between sets of concepts and the other concepts that are logically implied and correspondence rules that describe how syntactic structures are mapped to conceptual structures.

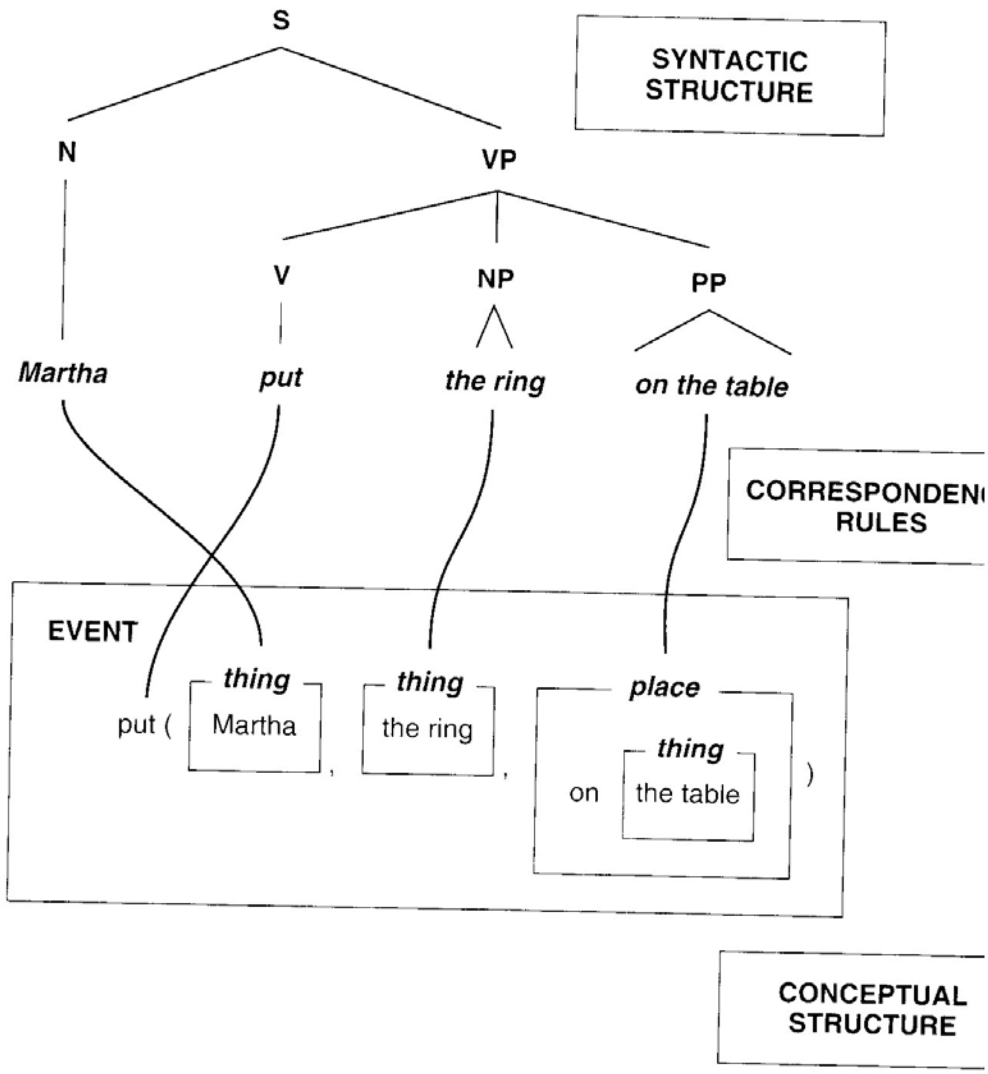

Figure 3. A schematic of Conceptual Semantics (Taken from Whitney 1998).

Finally, cognitive grammar (Lakoff 1987, Langacker 1987), like conceptual semantics, argues that meaning is a relationship between language symbols and mental concepts. However, cognitive grammar argues that the conceptual structure or idealized cognitive model (ICM, Lakoff 1987) is primary in the definition of both lexical and compositional semantics. ICMs are thought to be composed of basic level categories (e.g. chair, dog) and image schemas (e.g. container, path, part-whole). Understanding a sentence, then, is the process of composing a set of models into a coherent whole.

While the perspectives outlined above have formed the framework for how semantics has been conceived they are not computational in nature and cannot be employed in any straightforward manner in the construction of models of human operators. In particular, even cognitive grammar, which assumes that knowledge is learned as opposed to being innate, does not specify the nature of the learning mechanisms in a way that could be instantiated as a computer program. However, these theories do make distinctions that are important in understanding what learning algorithms must achieve. In particular, both conceptual semantics and cognitive grammar make a key distinction between lexical semantics and compositional semantics, which we will use to structure the remainder of this section.

2.2 Lexical Semantics

Early psychological work on lexical semantics focused on how conceptual knowledge is structured informed to a large extent by data on the speed with which simple statements such as ”A robin a bird.” could be verified (M. Collins & Quillian, 1969). While this work was useful in elucidating some of the general principles of semantic memory, from an applied perspective it was limited, as it provided no mechanism by which the representations of large numbers of words could be induced. Consequently, demonstrations were restricted to handcrafted toy examples in laboratory settings. More recently, however, interest in lexical semantics has shifted to mechanisms for acquiring word meaning from text corpora.

Interest in creating mechanisms for learning lexical semantics from text were ignited by provocative demonstrations of the Latent Semantic Analysis model (LSA, Landauer & Dumais, 1997). LSA constructs a term by document matrix of counts from a corpus, applies weighting functions and then takes the singular value decomposition of this matrix in order to remove noise and force similar words to be Human Memory represented by similar vectors. What was startling about the performance of the model was its performance on some tasks that require semantic knowledge. For instance, LSA was used to mark essay assignments by summing the vectors corresponding to the words in a student essay and comparing this against prescored student essays. The scores of the nearest neighbours are then averaged to get an overall score for the essay. Landauer and Dumais (1997) reported results in which LSA was able to perform as reliably as human tutors. Furthermore, LSA was applied to 80 retired items from the synonym component of the Test of English as a Foreign Language (TOEFL, Landauer & Dumais, 1997). In this task, applicants to US universities from non-English speaking countries choose from four alternatives the one that is closest in meaning to a stem word. To simulate this task with LSA, the cosine between the stem word and each of the alternatives is calculated and the alternative with the highest cosine is chosen. LSA was able to perform well enough to pass the test.

Initial models such Latent Semantic Analysis (LSA, Landauer & Dumais, 1997) and Hyperspace Analog to Language (HAL, Lund & Burgess, 1996) have now been supplemented by a number of models from the machine learning literature including Probabilistic Latent Semantic Indexing (Hofmann, 2001), Latent Dirichlet Allocation (LDA, Blei, Ng, & Jordan, 2002), the topics model (Griffiths & Steyvers, 2002), Word Association Space (Steyvers, Shiffrin, & Nelson, 2004), non-negative matrix factorization (Lee & Seung, 1999; Ge & Iwata, 2002), local linear embedding (Roweis & Saul, 2000), independent components analysis (ICA, Isbell & Viola, 1999), and the information bottleneck (Slonim & Tishby, 2000). It is beyond the scope of this review to describe each of these models in detail and as little comparative work has been done it is not yet clear which of these models performs best for which purposes. However, some of these models do have properties that are likely to make them preferable to LSA, at least in some circumstances. In particular, probabilistic methods like Probabilistic Latent Semantic Indexing, Latent Dirichlet Allocation and the topics model are built on firmer theoretical foundations. Furthermore, LDA, the topics model and ICA produce more interpretable representations and the topics model has been shown to do a better job of capturing the word association norms (Griffiths & Steyvers, 2002).

2.3 Compositional Semantics

Each of the perspectives on semantics outlined above, had at its heart a notion of compositional semantics, that is, how words are combined to make meaning. Unlike lexical semantics for which a reasonable attempt at meaning formation can be made by investigating the contexts of use of individual terms, the productive nature of syntax means that the majority of sentences will never have been seen before. Necessarily, the formation of the representation of a sentence must involve combining symbols in a systematic way that can be replicated by others to a degree sufficient to allow understanding. Furthermore, the simple additive notion employed in models like LSA will not suffice as clearly word order plays a role in compositional semantics. Models such as ACT (Anderson, 1983, 1993; Anderson & Lebiere, 1998) and Construction Integration theory (Kintsch, 1998) have emphasized the need for propositional representations to capture the relationships between concepts and preserve role assignment. In this section, we will review current attempts to automate the capture of propositional information from corpora (Dennis, Jurafsky, & Cer, 2003). The majority of the mechanisms that have been designed to address compositional semantics fall into the class of supervised semantic parsers. These models take as their task the generation of correspondence rules as outlined in conceptual semantics, that take a sentence and assign role labels (Agent, Patient, Instrument, Location, etc) to the relevant constituents for each of the predicates in the sentence (Blaheta & Charniak, 2000; Gildea & Jurafsky, 2002; O'Hara & Wiebe, 2002; Palmer, Rosenzweig, & Cotton, 2001). For instance, given the sentence:

Sampras outguns Agassi in Utah

a system would produce an annotation such as:

$\text{Sampras}_{\text{Agent}}$ outguns $\text{Agassi}_{\text{Patient}}$ in $\text{Utah}_{\text{Location}}$

This work has been driven, at least in part, by the availability of semantically labeled corpora such as Propbank (Kingsbury, Palmer, & Marcus, 2002) and FrameNet (Fillmore, Wooters, & Baker, 2001) which provide the relevant training data. For instance, the system proposed by Gildea and Jurafsky (2002); Pradhan, Hacioglu, Ward, Martin, and Jurafsky (2003), the sentence is first parsed, various syntactic, lexical, and semantic features are extracted, and then a classifier is applied. The classifier looks at each word chunk or parse constituent in the sentence, and decides for each one whether it is an argument of the verb, and if so what its semantic role is. These features include the syntactic type of the constituent (NP, PP, S, VP, etc), the identity of the verb, the path in the parse tree between the verb and the constituent, whether the constituent is before or after the verb, etc. Given enough training data, the classifiers achieve quite reasonable accuracy on semantic role labeling.

One important limitation of supervised semantic parsers is that they rely on the accuracy, coverage and labeling scheme of their training set. There are a great many schemes that have been proposed ranging in granularity from very broad, such as the two macro-role proposal of Van Valin (1993), through to theories that propose nine or ten roles, such as Fillmore (1971), to much more specific schemes that contain domain specific slots such as ORIG CITY, DEST CITY or DEPART TIME that are used in practical dialogue understanding systems (Stallard, 2000). It is well known to be difficult to label semantic roles in a general way, and each database (like FrameNet and PropBank) makes choices about roles that are often incompatible. Furthermore, the system's performance is quite dependent on the training data looking like the test data. Therefore, we think it is key to understand how this task could be extended to an unsupervised one.

The task addressed by the Syntagmatic Paradigmatic model (Dennis, 2004, 2005) is to create proposition-like units without recourse to labeled training data. The reasons for interest in an unsupervised method are two fold. Firstly, at a theoretical level the information that can be induced by the SP system forms a lower bound on what a human inductive system might achieve. People do not have access to semantically or syntactically labeled training data when learning a language and so in order for the system to be relevant to the debate on what can and cannot be induced it is crucial that the system be subject to the same constraints. Secondly, as indicated above, supervised semantic parsers tend to degrade significantly when applied to text that differs from that on which they were trained. Creating sufficiently large corpora of labeled training data is expensive, error prone and time consuming, however.

The key to the SP approach to unsupervised parsing is a switch from intentional to extensional representation. In systems that employ intentional semantics, as above, the meanings of representations are defined by their intended use and have no inherent substructure. The names of the roles are completely arbitrary and carry representational content only by virtue of the inference system in which they are embedded. Now contrast the above situation with an alternative extensional representation of Sampras outguns Agassi , in which roles are defined by enumerating exemplars, as follows:

Sampras: Kuerten, Hewitt
Agassi: Roddick, Costa

The winner role is represented by the distributed pattern of Kuerten and Hewitt, words that have been chosen because they are the names of people who have filled the X slot in a sentence like X outguns Y within the experience of the system. Similarly, Roddick and Costa are the names of people that have filled the Y slot in such a sentence and form a distributed representation of the loser role. The use of extensional semantics, of this kind, has a number of advantages. First, defining a mapping from raw sentences to extensional meaning representations is much easier than defining a mapping to intentional representations because it is now only necessary to align sentence exemplars from a corpus with the target sentence. The difficult task of either defining or inducing semantic roles is avoided. Second, because the role is now represented by a distributed pattern it is possible for a single role vector to simultaneously represent roles at different levels of granularity. The pattern Kuerten, Hewitt could be thought of as a proto-agent, an agent, a winner, and a winner of a tennis match simultaneously. The role vectors can be determined from a corpus during processing, and no commitment to an a priori level of role description is necessary. Third, extensional representations carry content by virtue of the other locations in the experience of the system where those symbols have occurred. That is, the systematic history of the comprehender grounds the representation.

For instance, we might expect systematic overlap between the winner role and person-who-is-wealthy role because some subset of Kuerten, Hewitt may also have occurred in an utterance such as X is wealthy. These contingencies occur as a natural consequence of the causality being described by the corpus and have been dubbed inference by coincidence (Dennis, 2004). To create extensional representations of sentences, Dennis (2004, 2005) used String Edit Theory (Sankoff & Kruskal, 1983) to align sentences from a corpus with the target sentence. As an illustrative example, suppose that we wish to create a representation for the sentence Sampras outguns Agassi in Utah and that the following are the most probably alignments across a corpus:

Sampras	outguns	Agassi	in Utah
Kuerten	downs	Andy Roddick	
Hewitt	defeats	Costa	

The extensional representation thus formed would be:

Sampras: Kuerten, Hewitt
Outguns: downs, defeats
Agassi: Roddick, Costa

To test the model, articles were taken from the Association of Tennis Professionals (ATP) website. Then questions of the form "Who won the match between X and Y? X" were created. Finally, the model was presented with the same sorts of questions with the answer omitted. An extensional representation of each question was created and matched against the extensional representations of the sentences from the target articles to identify sentences that might contain the critical information. From these sentences the filler most likely to be appropriate in the answer slot was chosen (see Dennis, 2004 for details). On 67% of occasions the model correctly returned the winner of the match. 26% of the time it incorrectly produced the loser of the match. 5% of the time it responded with a player other than either the winner or loser of the match and on 3% of occasions it committed a type error, responding with a word or punctuation symbol that was not a player s name. Interestingly, when the model was correct, on 41% of occasions the sentence that it chose as the most likely to contain relevant information was one from which an answer could be inferred but that did not contain a literal statement of the result. That is, inference by coincidence plays a significant role in the performance of the model. The model demonstrates that it is possible to extract proposition-like representations from open text using only simple string edit operations that have no built in grammatical or semantic knowledge. Furthermore, the importance of inference by coincidence in the model suggests that systems based on intentional representational systems may be throwing away a critical source of statistical information that may underpin the robustness of the human comprehension apparatus.

One objection to instance-based models, like the SP model, is that searching the number of exemplars that is necessary to realistically capture the relevant language structures is not feasible. However, there have been recent advances in methods for finding approximate nearest neighbours that provide a resolution to this issue. In particular, the method known as locality sensitive hashing (LSH, Marzal & Vidal, 1993; Gionis, Indyk, & Motwani, 1999) is now in wide spread use and shows significant promise. LSH relies on the observation that in most instance-based methods (including the SP model) it is not necessary to have the exact nearest neighbours of a stimulus in order to proceed- approximate nearest neighbours are sufficient. LSH creates hash functions which are designed to map stimuli into a hash table so that similar stimuli are likely to collide and be stored in the same cell. Finding near neighbours then involves applying the hash function and retrieving the stimuli that appear in the cell. For any given hash function there will be boundaries where similar stimuli are nonetheless mapped to different cells. LSH alleviates this problem by constructing a series of hash functions with randomly varied similarity functions. To find nearest neighbours all of the hash function are consulted and the joint set of the neighbours collected. As the number of hash functions increases it becomes increasingly less likely that a similar stimulus will fail to be retrieved. In this way, then, LSH can eliminate the difficulties associated with large instance sets.

Understandings of semantics and semantic memory have a long history. While semantic memory has been incorporated into unified models like Anderson's ATC-R theory (Anderson & Lebiere, 1998) that have been applied to human operator modeling, this involves hand crafting representations and rules. Such models generalize poorly and require constant maintenance in order to perform in new task environments. With the rise of corpus-based approaches to lexical semantics like LSA, the ability to perform practical tasks with automated methods has improved dramatically (see future work section for an elaboration of this point). Similarly, we suspect that it will be necessary to employ completely unsupervised mechanisms like the SP model before coverage will be sufficient to model compositional semantics and inferencing at a level that will be useful in human operator modeling.

3. Episodic Memory

Over the last 25 years the majority of models of episodic memory have fallen into the global matching category (Raaijmakers & Shiffrin, 1981; Gillund & Shiffrin, 1984; Murdock, 1982; Hintzman, 1984; Humphreys, Bain, & Pike, 1989). These models encountered difficulties with some data, in particular, the list strength effect (Ratcliff, Clark, & Shiffrin, 1990; Shiffrin, Ratcliff, & Clark, 1990) and have more recently been replaced by a set of Bayesian models (Dennis & Humphreys, 2001; Shiffrin & Steyvers, 1997; McClelland & Chappell, 1998). We will discuss this history, the current models, including network and connectionist models, and also discuss some recent work on memory access control processes. Issues about computational feasibility and brittleness have already been discussed in the semantic memory section.

3.1 Differentiating Episodic from Semantic Memory

In a free association test a subject is asked to produce the first word that comes to mind when presented with a cue word. Given these instructions approximately 50% of the population respond table when shown chair. Originally it was assumed that the learning of an arbitrary pair of words in the laboratory (e.g., chair pencil) would directly compete with the pre-existing association revealed in the free association task. However, it became evident that there was a very considerable separation between the laboratory acquired association and the pre-experimental association. In particular acquiring and association in the laboratory did not alter free association probabilities (Slamecka, 1966). This ability to keep the two kinds of memories distinct led Tulving (1976) to propose a distinction between episodic and semantic memory.

The episodic/semantic distinction has proved to be very popular thought it is still very poorly defined. Tulving's original conception was that episodic memory required a memory that was unique to an episode (the learning circumstances) whereas semantic memory was independent of any particular circumstances. Since Tulving (1983) there has been an emphasis on conscious awareness of the learning circumstances. There has also been a very strong claim that episodic and semantic memory constitute separate memory systems. Much of the support for the separate memory systems claim came from findings showing that task performance could be changed even when the task or test instructions made no reference to the learning episode. For example, exposure to a word within a long list of words will enhance the probability that the word will be produced given the first three letters as a cue or will be identified correctly if presented under degraded viewing conditions. These enhancements in performance occur for profoundly amnesic subjects as well as normal subjects and do not require any reference to the study episode. Further support for some form of separate systems comes from a computational analysis performed by (McClelland, McNaughton, & O'Reilly, 1995). They found that they could not capture two important aspects of human memory performance within the one model. One aspect involved the slow acquisition of the regularities and differences in the environment. The other aspect involved the very rapid acquisition of associations between arbitrary events or the knowledge that an item had occurred in a particular context. They found that models which acquired the regularities and differences in the environment produced far too much unlearning of previously acquired arbitrary associations. Likewise models that produced the rapid acquisition of new associations had considerable difficulty in generalizing to related stimuli.

Tulving (1976) and many researchers since then probably thought that a requirement for a unique memory meant that memories were stored separately (e.g., Flexser & Tulving, 1978). However, Humphreys, Bain, and Pike (1989) showed how in a composite memory associations with a contextual cue could serve to produce episodically unique memories. Humphreys, Wiles, and Dennis (1994) also argued that Tulving's (1976) distinction was useful because it acknowledged that different tasks had different goals. That is they proposed that the goal of the paired associate task was to produce the item which had been paired with the cue in the context specified by the instructions whereas the goal of the free association task while not being well specified was closer to producing the strongest associate of the cue.

In addition, to the lack of theoretical agreement there is some disagreement about classifying tasks as episodic or semantic. The paired associate and free recall tasks are almost universally regarded as episodic even though our understanding of how pre-existing semantic associations affect performance in both tasks is limited. There is, however, some dispute as to whether pair recognition requires an episodically unique memory (Humphreys, Bain, & Pike, 1989; Hockley, 1992) and there is some indication that pair recognition can be performed by someone who has severe damage to brain structures known to be important for other episodic tasks. There is also a dispute as to whether the more or less continuous information (familiarity) which seems to underlie much of single item recognition is episodic or not (Mandler, 1980; Humphreys, Bain, & Pike, 1989; Yonelinas, 2002).

Cued recall with an extralist cue also poses a problem for a strong separation of episodic and semantic memory. In this task the test cue has a pre-existing semantic relationship with the target but the cue and the target were not studied together in the study episode. There have been attempts to reduce cued recall with an extralist cue to a special case of paired associate learning. For example, Flexser and Tulving (1978) assumed that the target alone would be stored in episodic memory just as the cue and target are assumed to be stored together when they are studied together. Furthermore they assumed that retrieval was basically the same whether the cue had been studied with the target or the target was studied by itself. That is they assumed that it was the overlap in the encoding of the cue and the encoding stored in the memory trace that produced trace retrieval. In this conception it is the similarity in encoding between the memory trace table and the extralist cue chair which determines how effective a cue chair will be for table. Many contemporary researchers seem to believe in something like the Flexser and Tulving account or at least do not see any problem posed by extra list cues to ideas about separate memory systems. However, there has been a considerable amount of research, mostly by Doug Nelson and his colleagues, on cued recall with an extralist cue since Flexser and Tulving. At this time it appears very unlikely that a modified version of the Flexser and Tulving model would suffice for cued recall with an extralist cue.

3.2 Models

Over the last 25 years the majority of models of episodic memory have fallen into one of four camps. The most numerous camp have been the global matching Human Memory Included are SAM (Raaijmakers & Shiffrin, 1981; Gillund & Shiffrin, 1984), TODAM and CHARM which utilize the same basic mathematics (Murdock, 1982; Eich, 1982) Minerva II (Hintzman, 1984) and the Matrix model (Humphreys, Bain, & Pike, 1989). The Global Matching models are now being superseded by what we will refer to as the Bayesian models.

Included here is REM (Shiffrin & Steyvers, 1997), McClelland and Chappell (1998), and BCDMEM (Dennis & Humphreys, 2001). In addition there has been continuing work on network models in the Anderson and Bower (1973) tradition (Anderson, Bothell, Lebiere, & Matessa, 1998). Finally a series of connectionist models have been developed in order to account for priming effects. Because it seems increasingly likely that priming or implicit memory will have to be considered as part of any consideration of episodic memory we have also briefly reviewed these models.

TODAM and the Matrix model represent items as vectors and the association between items as either the convolution (TODAM) of two or more vectors or the tensor product of two or more vectors (Matrix model). The convolution and Tensor product operations are basically similar with the exception that the convolution operation is symmetrical whereas the tensor product operation is not. Storage and retrieval are conceptualized in these models in very similar ways and different retrieval operations are employed for recognition and recall. Storage starts with a psychological assumption about what is stored. Is it an item, an association between a context and item, an association between two items, or an association involving context and two items? We will illustrate the storage and retrieval process with reference to the situation where associations between two items are stored. In TODAM for each pair the convolution of the two item vectors would be created (this is also a vector) and the resulting vectors would be summed over every item in the list. Recognition is conceived of as a matching operation (a computation of the similarity between the test cue(s) and the composite memory. However, the mathematics permit two slightly different interpretations. If the task is to discriminate between intact and rearranged pairs in one interpretation one of the pair members would be used as a retrieval cue. That is, the vector representing that pair member is correlated with the composite memory vector. The result is a vector which contains the vector of the other pair member plus noise. At this point the dot product between the output vector and the vector representing the other pair member can be calculated. This is a measure of similarity between the input cues and the composite memory. In the other interpretation of the recognition process the convolution of the vectors in the test pair is calculated. The dot product between the convolution representing the test pair and the composite memory is then calculated. These are psychologically different interpretations but they are mathematically equivalent. As indicated in recall the cue (one member of the study pair) is correlated with the composite memory. Because the output vector is noisy it has generally been considered that some additional process is required in order to converge to a single response. A variety of procedures have been proposed which differ in their psychological plausibility. More complex memories are also possible in which item vectors and vectors representing the convolution of two or more items are stored in the same composite memory. These storage arrangements have proven useful in representing order information and in extracting information about a single item from a complex memory for several items. The later ability may provide an explanation for subjects ability to identify the missing item form a known set.

The Matrix model also utilizes a composite memory which can be conceptualized as a sum over vectors (an item memory), a sum over matrices (a pair memory) or a sum over tensors of rank 3 (e.g. a three-way memory involving a pair in a context). The recall/recognition distinction in the Matrix model is essentially the same as in TODAM and the same two psychological interpretations apply to the recognition process. In order to actually use a tensor product model to make a prediction about recall some means of selecting a response out of the noisy output is required. In order to accomplish this task Chappell and Humphreys (1994) added an auto associative memory. This is essentially an association between an item and itself where active units try to turn on other units which were also active during training and to turn off units which were not simultaneously active during training. Such a memory can remove noise converging to the vector with the largest weight in the output. However, such a memory does not always converge to a single item. Instead, at times, it can converge to a state where every element in memory is activated or to a state where no elements in memory are activated. Chappell and Humphreys (1994) interpreted the state of total of total activation and the state of no activation as corresponding to a situation where response competition had blocked the production of a response. Humphreys, Tehan, O'Shea, and Boland (2000) then provided experimental support for this idea. One implication is that with recall tasks relevant information may be in memory but no information relevant to the task may be available. In contrast, recognition (matching) tasks always produce relevant information. In another extension of the Matrix model Wiles, Humphreys, Bain, and Dennis (1991 ; also see Chappell and Humphreys, 1994) explored a variety of ways to compute the intersection between two sets of activated items. The thinking here was to retrieve an item which was in both sets (e.g., an item which was activated by a contextual or list cue and an item which was activated by a semantic cue in order to provide a model for cued recall with an extralist associate cue. The idea of an intersection was then further extended by Humphreys et al. (1994) who argued that when recognition is conceived of as retrieving using one cue and matching the output with another cue is essentially the same as tasking an intersection. This thinking also extends to the confirmation of an expectation where an animal may anticipate a particular sensor input and match the incoming stimulus against the anticipated input.

Minerva II and SAM assume that items are stored separately. Minerva II uses a vector representation where an item is represented by a vector and a pair of items is represented by concatenating the two vectors representing the individual items. Storage consists of placing a vector representing an item or a SAM represents an item as a symbol with no internal structure. These are referred to as images and what is stored is the association between a word and an image or between a context and an image. Minerva II constructs a vector representation of a test probe and matches this probe (takes the dot product) against each memory. The match value is then raised to the third power and all of the strengths of these individual matches are summed. In SAM the product of the strength of the association between each cue and an image is computed and these products are then summed over every image in memory. In spite of these differences Minerva II SAM, TODAM, and the Matrix model are highly similar for recognition. In fact, Humphreys, Pike, Bain, and Tehan (1989) showed that they could all be subsumed under a more general model. This occurs because the match with the composite memories (TODAM and the Matrix model) can be written as a sum over the matches with the individual memories that were added together to make up the composite and because the product of associative strengths in SAM acts very much like a match between the cues and the individual memories.

Cued recall in Minerva II is accomplished by retrieving a composite vector. To create the composite each vector in memory is weighted by the cube of the match between the probe vector and the memory vector. Then the composite vector is fed back to the memory system as a new cue. After several iterations of this process Minerva II tends to converge on the vector in the composite which initially had the strongest weight. Of course a coalition of similar vectors in memory can produce a result which is the central tendency of the set of similar vectors. For recall SAM used a more conventional idea. The idea here is that images in memory were sampled with a probability which depended on the product of the associative strengths between the cues and the image.

It was noted that all of the Global Matching models predicted a list length and list strength effect (Ratcliff et al., 1990; Shiffrin et al., 1990). That is in all four models the learning of additional items increases the matching strength of targets and distracters by the same amount. It also increases the matching strength variance so the overall ability to discriminate between targets and distracters declines. In the three models which use vector representations this increase in the variance of the matching strengths is an inevitable consequence of the vector representation. In the SAM model the increase in variance was built into the model from the beginning though it is not an inevitable consequence of the representation assumptions. The strengthening of some items in a list via repeated presentations or extra study time also increases matching strengths for non-strengthened targets, for distracters and their variances. Tests of these predictions showed little or no list strength effect in recognition. At the same time there appeared to be a list length effect. SAM and presumably Minerva II could be modified in order to accommodate these findings. In order to do this it was assumed that repeated presentations increased differentiation. In order to implement this in SAM it was assumed that the average strength of the association with other list items decreased with repeated presentations or more study time.

Differentiation was not a possibility with composite memories so composite memories appeared unviable. However, it was shown that if pre-existing memories were included in TODAM that the magnitude of the list strength effect could be reduced to what appeared to be an acceptable level. In addition, Dennis and Humphreys (2001) challenged the assumption that robust list length effects existed in recognition. Their argument was that list length effects were quite small when appropriate controls were used such as controls for retention interval, for contextual reinstatement, for fatigue, and for displaced rehearsals. At this time it appears that the absence or near absence of a list strength effect in recognition memory is not an impediment to the existence of composite memories. Nevertheless the list strength finding was one of the main motivators to the new generation of memory models.

3.2.1 Bayesian Models

Another motivator for the new generation of models was the ubiquity of the mirror effect in recognition. With many material variables the items which are learned more readily (e.g. low frequency words) have both a higher hit rate and a lower false alarm rate. This effect could be handled by the Global Matching models only by making a series of arbitrary assumptions. The new generation of models sought a less arbitrary explanation for the mirror effect. REM (Shiffrin & Steyvers, 1997) was a successor to SAM. It represents words as vectors and includes both a semantic and episodic memory. When a representation of a word is stored in episodic memory the features of that representation are supplied from the representation in semantic memory. With longer study time or additional presentations more features are filled in which produces the enhanced differentiation. In realistic versions of REM context features as well as item features are stored in semantic memory. At test the probe vector which contains context features and the item features from the probe is matched in parallel against the episodic images of all of the words on the list. Various versions of REM have used slightly different assumptions about how the context features are used. The simplest assumption is that they are used to activate images in memory and only images whose activation exceeds a threshold are included in the subsequent calculations. After the activated set of images has been selected the corresponding positions in the probe vector and each image in memory are compared to see if they are the same or different. This comparison ignores positions where no value is stored in the episodic image. There is also no contribution to the calculation resulting from matches between contextual features in the probe and in the memory images because these matches do not differentiate targets and distracters. The probability that the pattern of matching and mismatching positions would have occurred if the image was derived from the semantic representation of the target divided by the probability that it would have occurred if the image was not derived from the semantic representation of the target is then calculated. This calculation takes into consideration the relative likelihood of different features being present (that is properties of the semantic memory) but does not take into consideration properties of the learning situation (e.g., the proportion of low and high frequency items in the study list). However, it is acknowledged that some features of the learning situation which are obvious such as the list length or the number of study presentations might be taken into consideration in setting the criterion for responding. Finally Bayes rule is used to calculate the probability that the test word is represented in the set of activated images. With the assumption that likelihood odds are being calculated the mirror effect falls out naturally form most manipulations of learning diffculty. For example, Shiffrin and Steyvers (1997) proposed that high frequency words were less well recognized because they consisted of features which had a higher probability of occurrence. REM has been applied to an impressive array of paradigms including priming effects in semantic memory.

McClelland and Chappell (1998) in several aspects is similar to REM. Like REM it assumes that items become more differentiated with more learning. It also does this by storing separate representations of items and by learning about the features that make up those representations. However, the learning mechanism is sufficiently different for it to be considered as an alternative. Very briefly item detectors in M&C learn about the probabilities with which features occur. That is, as each item is presented only a subset of its features are activated. Item detectors have an initial hypothesis about which features will occur and this hypothesis is updated by which features actually occur. Over study trials the item detectors gain a better and better estimate of the features which occur when a given item is presented. At test the test items are compared to each item detector and the probability that the test item is the stored item is calculated.

Dennis and Humphreys (2001) also responded to the challenge posed by the finding of a null list strength effect. Their response was to assume that while recall was an item noise process recognition was a context noise process. That is the Global Matching models and REM had assumed that the primary source of noise was the other items in the list. This was very explicit in models which either did not incorporate context or which assumed that context isolated the list items from other memories so that only list items contributed to noise. In contrast Denis and Humphreys proposed a model (BCDMEM) where subjects reinstated a context and then matched the reinstated context against a context which was retrieved using the test item as a cue. Like REM and the McClelland and Chappell (1998) model BCDMEM used a likelihood odds calculation as an alternative to a direct comparison of the similarity of the reinstated and retrieved contexts. This enabled them to predict a word frequency mirror effect by making reference only to an empirically observable difference between high and low frequency words. Namely that high frequency words occur in more contexts. There is also increasing evidence that the impact of the other list items is less than had been assumed in the Global matching models (Maguire, Humphreys, & Dennis, in preparation). Whether the ultimate resolution of this issue will favour Dennis and Humphreys' extreme proposal (no item noise in recognition) or Criss and Shiffrin's (2004) compromise proposal (both item noise and context noise effects) remains to be determined.

3.2.2 Network Models

Anderson's ACT-R model (Anderson & Lebiere, 1998; Anderson et al., 1998) incorporates a spreading activation implementation of memory. This memory system is coupled with a production system that regulates the flow of information between long-term memory, working memory and a set of perceptual and motor buffers. Within this system, retrieval from memory is a serial process, stochastically selecting elements from the network with a high level of activation that match the current retrieval cue. Nodes in ACT-R's semantic network have a base level of activation that reflects both the frequency and recency of previous retrievals. The activation of a node decays over time, but receives a boost in activation each time it is accessed. Furthermore, frequent retrievals effect the decay rate of the concept node, allowing frequently used concepts to have an overall higher base level of activation. Thus, ACT-R makes explicit assumptions as to the amount of activation present, item repetition, and varying levels of general word frequency. Retrieving information from memory in ACT-R uses a parallel matching, serial retrieval mechanism in which all the elements in memory are compared against the request, with only a single item being retrieved. In selecting the item, temporary activation spread from the cue is added to the base level of activation of each node, with the most active element from memory being retrieved. The overall activation of a node is determined by four main factors. The first factor is the node's base level of activation. As explained above, this will favour the retrieval of recently processed items (demonstrating priming and repetition effects), or items that are generally more frequent (accounting for word frequency effects).

Secondly, ACT-R's architecture contains a form of working memory (called the "goal buffer") that stores information that is being attended to. An amount of attentional activation is divided evenly amongst the nodes corresponding to these attended concepts, and is allowed to spread to their neighbouring nodes within the network. The third factor in determining the activation of a node during retrieval enhances the activity of nodes matching the actual retrieval request, and suppresses nodes that do not. This differs from the second factor that represents contextual influences upon recall, activating all concepts that are connected to anything under the attentional spotlight. As there is a threshold that determines whether or not a retrieval has been successful, this factor enhances the likelihood that this threshold will be met by items matching all the criteria (such as the fact that it is a WORD that is associated with DOG that was on the study list etc.). The last term added in the calculation of the activation is noise associated with storage and retrieval of information. As ACT-R always retrieves the item with the most activation, gaussian noise is added so that the item retrieved is not deterministic, but rather, is probabilistically biased.

In contrast to network models such as ACT-R in which concepts are represented by distinct nodes, connectionist networks have also been frequently employed to model aspects of human memory. These models differ from ACT-R in that concepts are represented as overlapping distributed vectors of features, with nodes often appearing in distinct processing layers (e.g., a letter feature level, a letter level and a word level in the case of modelling word reading). In connectionist networks, information flows between units over time, but at a subsymbolic level (as no single connection can be viewed as a direct connection between distinct concepts). To learn the appropriate network weights to perform a given mapping between network input and output, powerful learning algorithms such as backpropagation are employed (see Haykin, 1994). Frequently, if the mappings are non-linear, "hidden" intermediate layers are used, with the system being required to automatically learn the appropriate representations. As with LSA, neural networks can be trained on raw data (such as text corpora), to form representations analogous to those believed to exist in semantic memory. For example, Elman (1991) trained a simple recurrent network (that affords the processing of temporal sequences) to predict the probability distribution for possible next words of the sentence, given a word at a time as input. Although the word representations at input and output used a local coding scheme (with each word being represented by a different node), the hidden representations that were learned over time reflected the meanings of the words, with similar words and word categories yielding similar representations. Similar networks have also been used as an alternative approach to LSA for representing paragraphs of text as distributed vectors of latent features that afford the retrieval of similar information (O'Reilly & Munakata, 2000).

Unlike models such as ACT-R that attempt to model semantic and episodic memory within a unitary architecture, different aspects of memory are generally modelled by different neural network models and mechanisms. As mentioned earlier, McClelland et al. (1995) suggest that following a connectionist approach, semantic and episodic memory should be modelled as separate systems as they have conflicting requirements. That is, episodic memory requires fast learning, which in a neural network requires non-overlapping and sparse representations, whereas, semantic memory requires a slow learning rate and overlapping distributed representations as its general aim is to extract the statistical structure of the environment. Short term memory (or working memory), similarly can be described in terms of different requirements, being viewed as a maintained state of activation rather than a change of connection strength (O'Reilly & Munakata, 2000).

Apart from modelling the acquisition of semantic information representations from experience, connectionist networks are well suited for modelling the retrieval processes involved in memory. For example, recurrent neural networks can perform pattern completion (i.e. retrieving associated knowledge given a cue), by falling into a learned stable pattern of activation (i.e. an attractor) over time (Hopfield, 1982; Ackley, Hinton, & Sejnowski, 1985; Seung, 1998). Such networks have also been used to model aspects of implicit memory. For example, in the semantic priming paradigm, words are read faster and more accurately when preceded by a semantically related item compared to an unrelated item (see Neely, 1991, for a review). Such effects can be explained either in terms of the overlap in semantic representations (with the network having less distance to travel to fall into the new attractor), or in terms of biases in the recurrent weights that support the transition between items that frequently co-occur in the environment (Plaut, 1995; Cree, McRae, & McNorgan, 1999).

In terms of modelling episodic memory using the connectionist approach, the most prominent exemplar is the (O'Reilly, Norman, & McClelland, 1998) hippocampal model. This model, apart from modelling the anatomy and physiology of the hippocampus, captures many central properties of episodic memory, such as the rarity of false recollection and the decrease in recollection but not recollection quality (the accuracy of retrieved information) with increased interference. In accordance with the constraints mentioned above, this model allows fast learning of episodes by representing them in terms of large, sparse and pattern-separated representations. These representations are formed through random wiring between the inputs (the individual components of the episode) and a much larger "hidden" layer that represents the whole episode (i.e. a conjunction of the inputs). Representation is reciprocally connected to a number of other layers (through fast learning connections), that form a large attractor network. Given partial information, as with attractor networks of semantic memory, the model can reinstantiate the whole episode and the individual bound components. One drawback of this model however, is its overall complexity compared to other models of memory, making much of the processing opaque and unintuitive.

3.3 Comparing the Different Kinds of Models

The models differ in their psychological assumptions as well as in their assumptions about the underlying structure of human memory. They also differ in the range of paradigms to which they have been applied. Some theorists have also taken as their primary goal the obtaining of detailed fits of empirical data where others have been content to use a model to illuminate the underlying psychological assumptions. It is clear that models with very different representational and processing assumptions can provide very good fits to the same data. It thus seems probable that assumptions about the cues used and what is bound to what (e.g., the issue of interitem or positional associations in serial learning) are more important than the assumptions about the underlying structure of memory. At this stage they probably have three main weaknesses. First even the most recent ones such as REM only capture a bit of the complexity of the relationship between semantic and episodic memory. Second, with a bit of an exception for ACT-R the control processes are external to the models. That is, it is the experimenter not the model which decides to retrieve an episodic memory as opposed to a semantic memory. Finally, the models have been primarily if not exclusively applied to laboratory tasks where the external world has been partitioned into events and episodes (e.g., the presentation of a word in an experiment). It will be a major challenge to apply ideas about episodic memory to the problem of retracing the path you took to climb a mountain.

3.4 Controlling Memory Access

During the 1990s a substantial amount of empirical research was based on ideas about recollection and familiarity. These concepts were used in a variety of different ways and are not well defined. However, the prevailing assumption was that what was retrieved from memory did not depend on what the person was trying to retrieve. That is, a feeling of familiarity was automatically produced when a subject read or heard a word and this did not depend on whether the person was trying to identify the words which occurred in list 2 or to identify the words which had been presented auditorially. Unlike familiarity recollection was considered to be strategic. However, the recollection process simply returned symbolic information which the person could inspect or search in order to come up with an appropriate decision. The Dennis and Humphreys proposal that subjects reinstate a context and then match the reinstated context against a retrieved context allows the question which is posed (was the word presented auditorially) to influence the output of the memory process. Evidence for this specificity of the memory retrieval process has now been produced (Humphreys et al., 2003; Humphreys, Weeks, & Hockley, n.d.). In addition, Jacoby and his colleagues have used a three stage process to also produce evidence for a degree of control over the memory access process. In the first stage subjects deeply process the study items in one list and shallowly process the study items in another list. They then receive a recognition test on one of the two lists. At the start of this test they are fully informed about the manner in which the old items were processed. For example, they are told that all of the old items on the test list were rated for pleasantness or that all of the old items on the test list were rehearsed. In the third stage they receive a second test list for the lures on the first test list. In this test they are supposed to say yes to any word which appeared on the first test list and no to words which occur for the first time on the second test list. The results show that memory for test lures depends on the subjects expectations about how the old items were processed. Conceptually similar results were obtained by Maguire et al. (in preparation). They had subjects study two lists where half of the words in each list were presented visually and half auditorially. At test half of the subjects were instructed to say yes to list 2 words and no to list 1 words and to new words. The other half were instructed to say yes to read words and no to heard words and new words. Study words were randomly assigned to be in list 1 or 2 and to be presented auditorially or visually. A fixed set of 80 new words were used. There were 20 high frequency high contextual variability words, 20 high frequency low contextual variability, 20 low frequency high contextual variability words and 20 low frequency low contextual variability words. There was a main effect of frequency and a main effect of contextual variability on false alarm rates. In addition there was a test instruction by frequency interaction. For low frequency words there was no difference in the FARs between those subjects who were asked to say yes to list 2 words and those who were asked to say yes to read words. However, the subjects who were asked to say yes to the read words had higher false alarms on the high frequency words than did the subjects who were asked to say yes to the low frequency words. What these examples have in common is that they all suggest that test items are processed differentially depending on the question posed to the memory system. It seems probable that this kind of control process will receive further attention.

4. Prospective Memory

In recent years, memory researchers have shown increasing interest in memory tasks where an individual must remember to perform an action at some designated point in the future. The term prospective memory has been used to refer both to the tasks used and to a hypothetical type of memory which is assumed to underlie performance on the tasks. We will use the term prospective memory but we do not assume that there is a substantial separation between the processes employed in these tasks and the processes employed in the more traditional episodic or retrospective memory tasks. First, this section defines prospective memory and describes how it is assumed to differ from the more extensively studied retrospective memory tasks. We then introduce two theoretical frameworks that have been used to explain prospective memory performance, and examine evidence that supports these approaches. We finish by introducing an alternative prospective memory laboratory paradigm, which examines the execution of intended actions that have been briefly delayed. This type of prospective memory task is particular relevant to operators with many competing task demands.

There are two main types of prospective memory tasks. Event-based prospective memory tasks require individuals to remember to perform an action when a particular target event occurs in the environment. Examples include remembering to stop at the post office when driving past, or remembering to give your colleague a message next time you seen him. In contrast, time-based prospective memory tasks require actions to be performed at certain points in time (Einstein & McDaniel, 1990). The majority of research to date has concerned event-based prospective memory, and it is this type of prospective memory that forms the focus of the review in this section.

In everyday event-based prospective memory situations, individuals are often busily engaged in other activities in the time interval between planning an action and the time that an environmental target is encountered. In order to execute the delayed intention, individuals must interrupt these ongoing activities. Similarly, laboratory-based prospective memory tasks typically require participants to perform a special action (e.g., press the F1 Key) upon presentation of a specific event (e.g., the word dog) whilst performing an unrelated ongoing activity (e.g., rating words) (Einstein & McDaniel, 1990). The defining feature of event-based prospective memory tasks is that there are no external agents (e.g., experimenter, printed instructions) directing participants to engage in a memory search. It has been generally assumed that this differentiates prospective and retrospective memory tasks. However, this difference is most likely a matter of degree. That is, instructions are frequently given at the start of a retrospective memory test (e.g., a free recall trial or a list of words to recognize) and must be maintained while performing that test. Nevertheless, prospective memory tasks have a much greater emphasis on the maintenance of an intention to remember and or place a heavy emphasis on the ability of the cue to initiate an intent to remember as well as the recall of the to-be-performed action.

In a typical laboratory event-based prospective memory task, participants engage in an ongoing activity. For example, participants might rate the pleasantness of words, make lexical decisions, or judge the frequency of words. In addition, when certain targets (or a class of targets) appear in the context of the ongoing task, participants are required to remember to make some overt response (e.g., press the F1 key) to indicate that they have remembered the intention. Much is already known about the factors that influence whether individuals will remember to perform an intention. For example, targets that are particularly salient (Einstein, McDaniel, Manzi, Cochran, & Baker, 2000; Marsh, Hicks, & Hancock, 2000), distinct from background context (McDaniel & Einstein, 1993), or highly associated with the intended action (McDaniel, Einstein, & Breneiser, 2004) tend to be noticed and responded to more frequently. Under conditions where individuals are required to remember to respond to a specified category of targets, prospective memory performance increases when highly typical category members are presented as targets (Ellis & Milne, 1996). Not surprisingly, prospective memory performance also improves when the prospective memory task is perceived to be more important than the ongoing task (Kliegal, Martin, McDaniel, & Einstein, 2004).

4.1 Theoretical Approaches

One prospective memory theory is that the processes involved are relatively automatic (e.g., Einstein & McDaniel, 1996; Guynn, McDaniel, & Einstein, 2001; McDaniel, 1995). That is, it is assumed that the process is driven from the cue. The alternative view is that successful prospective memory performance requires the allocation of resources in the period before the presentation (e.g., Burgess & Shallice, 1997; Guynn, 2003; R. E. Smith, 2003; R. E. Smith & Bayen, 2004). Thus this approach has an emphasis on identifying control processes and the mechanisms which maintain them. We will now focus on two prominent theories of event-based prospective memory, and report data that supports the mechanisms that these theories.

According to the Preparatory Attentional and Memory Processes (PAM) theory of prospective memory (R. E. Smith, 2003; R. E. Smith & Bayen, 2004), prospective memory requires cognitive resources. PAM theory distinguishes between the prospective and retrospective components of prospective memory. The prospective component is proposed to involve the allocation of resource-demanding preparatory attentional processes to monitor the environment for the presence of target events and the opportunity to perform intended actions. These processes are thought to operate prior to individuals attending targets. Furthermore, PAM theory claims that preparatory attention is functionally related to prospective memory performance, such that the amount of preparatory attention directed toward the prospective memory task will be positively related to prospective memory performance. Other researchers have proposed similar monitoring systems that serve the function of maintaining the cognitive system in retrieval mode (Guynn, 2003), maintaining increased levels of activation of targets through rehearsal (Guynn, 2003; McDaniel & Einstein, 2000), or interrupting ongoing activities when targets are attended (Burgess & Shallice, 1997). Upon actual presentation of target events, PAM theory proposes that retrospective memory processes are required to recognize targets and retrieve intentions.

Research has provided support for role of preparatory attention in prospective memory. First, findings that the addition of a demanding secondary task reduces prospective memory performance (e.g., Marsh & Hicks, 1998) support the general notion that the detection of target events requires cognitive resources. Second, prospective memory tasks embedded in ongoing tasks have been shown to slow performance on ongoing tasks, even at times when target events were not being presented (Guynn, 2003; Marsh, Hicks, & Cook, 2005; Marsh, Hicks, Cook, Hansen, & Pallos, 2003; R. E. Smith, 2003). For example, R. E. Smith (2003) found that participants with an event-based intention (i.e., remember to press the F1 key if studied words are presented) took 200-300ms longer to make lexical decisions on non-target trials compared to participants only performing lexical decisions. Third, response costs to non-target ongoing task trials have been found to positively correlate with prospective memory performance (R. E. Smith, 2003; R. E. Smith & Bayen, 2004), and response costs on trials preceding prospective memory hits have been found to be larger than on trials preceding prospective memory misses (West, Krompinger, & Bowry, in press), signifying a functional relationship between preparatory attention and prospective memory.

Other evidence on the nature of the control processes comes from an examination of the relationship between the prospective memory task and the ongoing activity. For example, Meier and Graf (2000) found that prospective memory performance was higher when the ongoing task and the prospective memory task required the same kind of processing (i.e., perceptual-perceptual, semantic-semantic) rather than different kinds of processing (i.e. perceptual-semantic, semantic-perceptual). Similarly, (Marsh et al., 2005) demonstrated that individuals were more likely to likely to remember to press the enter key when presented with palindrome targets (words that could be spelt the same way backwards and forwards) if these palindromes were embedded in an ongoing task that was orthographic in nature (decide if words have double letters), compared to an ongoing task that was semantic in nature (lexical decision).

An alternative theoretical approach to prospective memory claims that the processing of prospective memory targets can automatically bring to mind intended actions. This approach is known as the reflexive-associative processes view (Guynn et al., 2001; McDaniel et al., 2004). This alternative to PAM theory was inspired by the observation that event-based prospective memory bears striking similarity to cued recall (McDaniel & Einstein, 1993). Similar to the Moscovitch (1992) automatic-associative memory subsystem, the view claims that if attention to targets interacts sufficiently with the representation of intended actions, intended actions will be automatically delivered to awareness. In support of this view, there are certain conditions (e.g., when targets are salient) under which older adults perform as well as younger adults in event-based prospective memory (e.g., Einstein & McDaniel, 1990), suggesting that prospective memory performance under these conditions is controlled by automatic-associative processes and requires few cognitive resources.

According to the multiprocess view of prospective memory (McDaniel & Einstein, 2000; McDaniel et al., 2004), both non-automatic and automatic processes play a role in prospective memory retrieval, and the relative contribution of each depends upon task conditions. More specifically, the multiprocess view claims that prospective memory retrieval is more likely to be automatic when intended responses are simple, target events are distinctive, the ongoing task focuses processing on the relevant dimensions of targets, or target events are highly associated with intended responses. One particular variable receiving recent attention is the degree of association between targets and responses. According to the multiprocess view, the processing of target events that are highly associated with intentions are more likely to lead to the reflexive or obligatory retrieval of intended actions (e.g., reflexive-associative processes; Guynn et al., 2001; McDaniel et al., 2004). In contrast, when target events are not associated with the intended actions, processing of target events is less likely to lead to the automatic retrieval of intentions, and successful prospective memory performance is dependent on the strategic processes required to notice the significance of the target events and retrieve associated intentions (labeled cue-focused processes; McDaniel et al., 2004).

Recent research has provided support for the automaticity of prospective memory retrieval when prospective memory targets are highly associated with intentions. In a study by McDaniel et al. (2004), participants were required to remember to write down specific words when presented with target words during a word rating task. In the high association condition, target words were highly associated with response words (e.g., spaghetti-sauce), whereas in the low association condition target and response words were not associated (e.g., thread-sauce). The addition of a secondary digit-monitoring task reduced the prospective memory performance of the low association but not the high association condition. Furthermore, pre-exposure of non-target words (thereby making them less distinctive from target words) reduced the prospective memory performance of the low association but not the high association condition. Employing similar task conditions, Marsh et al. (2003) found that response costs to an ongoing task on prospective memory target trials were significantly larger under conditions of low association than under conditions of high association, suggesting that the retrieval of intended actions under conditions of high association required fewer cognitive resources.

4.2 Delayed-Execute Prospective Memory

A characteristic of standard prospective memory paradigms is that the participants are allowed to perform the action immediately after the presentation of targets. In many respects, this is representative of real world situations. For example, a police officer may intend to write down the number plate of an abandoned car upon arrival at a crime scene, and upon seeing the car, if he/she remembers to do it, can immediately perform this action. McDaniel, Einstein, Stout, and Morgan (2003) recently noted that the standard prospective memory paradigm fails to capture a common element of prospective memory tasks. In real world settings, execution of an intended action must often be briefly delayed; perhaps because the conditions for performing it are not appropriate, the person has been disrupted by another higher priority task, or the current workload in too demanding to respond immediately. For example, upon arriving at the crime scene, the police officer may remember that the number plate needs to be recorded, but has to delay this task until a response has been made to a new radio call that has just been received. Delays and interruptions of this type seem to be prominent in demanding work settings where forgetting can have serious consequences for safety, such as in aviation operations (Loft, Humphreys, & Neal, 2003). McDaniel et al. (2003) labeled this prospective memory situation delayed-execute prospective memory.

Researchers have only recently begun to examine the cognitive processes that support the maintenance and execution of intentions over brief intervals. The handful of publications in this area has focused on task conditions that influence the likelihood that, once an intention has been retrieved, individuals will remember to perform intended actions after a brief interval (Einstein, McDaniel, Williford, Pagan, & Dismukes, 2003; McDaniel, Einstein, Graham, & Rall, in press; McDaniel et al., 2003). The laboratory paradigm developed by McDaniel et al. (2003) requires participants to complete a series of one-minute tasks for about half an hour. At the outset of the experiment, participants are instructed to press a designated key on the keyboard whenever they see a red screen, but not until they have finished the current ongoing task. A salient (i.e., red screen) is used as the prospective memory cue to ensure that all participants retrieve intentions. This paradigm has been used to investigate the effects of (a) the length of the delay between the offset on the red target screen and the occurrence of the task change (e.g., 5 secs, 15 secs, 40 secs), (b) an increase in attentional demands caused by additional tasks, and (c) a brief interruption (e.g., 15 secs) via the introduction of a new task during the delay period.

When the ongoing task is not particularly demanding, participants do fairly well at remembering to perform delayed actions (e.g., 92% success rate found by Einstein et al., 2003). Nevertheless, maintaining intentions over brief delays poses diffculties for the human cognitive system, particularly when extra concurrent tasks are being performed ((e.g., McDaniel et al., 2003) or when individuals are interrupted by a new task during the delay period (e.g., McDaniel et al., in press). A theoretical model has been developed to account for how people remember to self-initiate the performance of intentions after brief delays. According to the active-maintenance account of delayed-execute prospective memory, individuals periodically activate their intention for performing the action over the delay period (Einstein et al., 2003; McDaniel et al., in press, 2003). According to this view, when participants encounter delays they periodically bring intentions into focal awareness. This process is presumed to be supported by either the associative relation between the intentions and cover activities (contextual cueing), or the strategic checking of uncompleted intentions (c.f. Ellis, 1996). Several important findings have emerged to support this theory, as reviewed below.

Firstly, in contrast to the classic forgetting function found in retrospective memory tasks, delayed-execute prospective memory performance does not decline with the length of delays (Einstein et al., 2003; McDaniel et al., in press). This result would be expected if individuals were periodically refreshing their intentions to perform an uncompleted action. Secondly, dividing attention with a digit-monitoring task negatively reduces the probability that the intended action will be completed at the end of the delay period (Einstein et al., 2003; McDaniel et al., 2003). This finding is consistent with the active-maintenance view, as increasing the demands of cover activities should interfere with periodic activation processes. Thirdly, brief interruptions of 10-20 seconds, caused by the introduction of new tasks during the delay periods have a negative effect on the maintenance and execution of intentions (Einstein et al., 2003; McDaniel et al., in press). According to the active-maintenance view, interruptions, by virtue of switching attention to a new task, should make it difficult to periodically retrieve intentions because they take attention away from the ongoing task (including the prospective memory task). Finally, (Einstein et al., 2003) found that the accuracy and speed on performance on ongoing tasks declines during delay periods, suggesting that participants use cognitive resources to periodically refresh incomplete intentions.

5. Categorization

Categorization in humans is mediated by multiple and qualitatively different cognitive pathways depending on the task being performed. According to Ashby and Maddox (2005), the various tasks that yield different cognitive, neuropsychological and neuroimaging results can be classified into four main categories: rule based tasks, information-integration tasks, prototype distortion tasks, and weather prediction tasks. There is no single model of categorization that accounts for performance on all tasks. Below is a review of the various forms of categorization task and what cognitive pathways are thought to be employed. This review is followed by an overview of the prominent categorization models and the categorization tasks to which they best describe human performance.

5.1 Category Learning Tasks

Rule-based Tasks: In rule-based category learning tasks, the rule that maximizes segregation between categories is one that is easy for subjects to verbalise (Ashby, Alfonso-Reese, Turken, & Waldon, 1998). That is, in this experimental paradigm there exists some conceptual label for the property values upon which the segregation is based (such as discriminating between objects based upon standard colours). An example task commonly used in neuropsychological assessment is the Wisconsin Card Sorting Test (WSCT), in which subjects must learn the rule that separates piles of cards containing geometric figures that differ in colour, shape and number (Heaton, 1981).

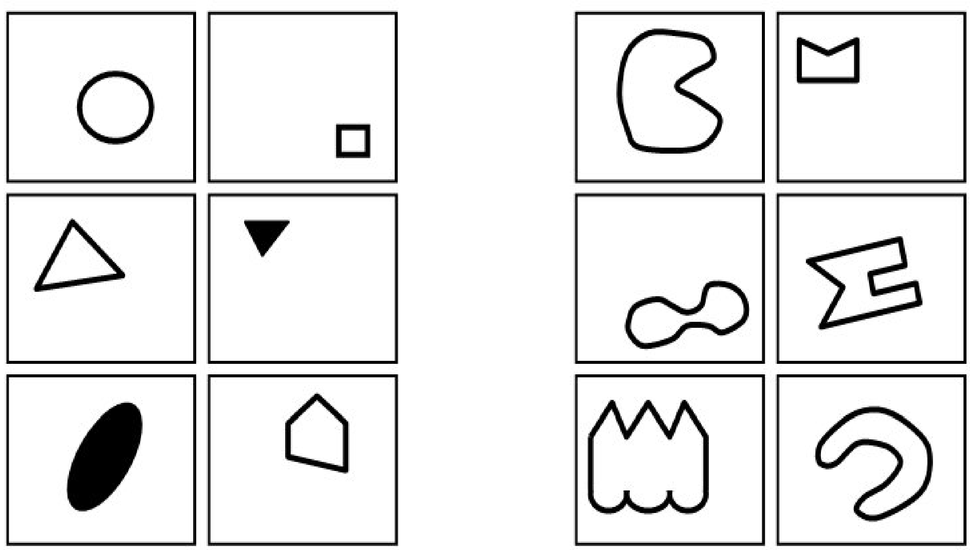

Figure 4. An example Bongard problem (redrawn from Bongard, 1970, p215). The aim of the task is to find a rule which segregates the six figures to the left from the six figures to the right.

Both neuropsychological and neuroimaging data support the view that humans solve rule-based tasks through the use of explicit reasoning processes. For example, in tasks similar to the WCST, task related activation was found in the prefrontal cortex and in the head of the caudate nucleus (Konishi et al., 1999; Lombardi, Andreason, Sirocco, Rio, & Gross, 1999; Rao, Bobholz, Hammeke, Tosen, & Woodley, 1997; Rogers, Andrews, Grasby, Brooks, & Robbins, 2000; Volz, Gaser, Rzanny, & Mentzel, 1997), with these regions being implicated in the use of executive attention (Posner & Petersen, 1990) and working memory (Goldman-Rakic, 1987, 1995). Consistent with these findings is the result that performance at these tasks is degraded in patients with frontal lobe damage (Kimberg, D’Esposito, & Farah, 1997) and Parkinson’s disease (where there exist deficits to the caudate nucleus (Ashby, Noble, Filoteo, Waldron, & Ell, 2003; R. G. Brown & Marsden, 1988).

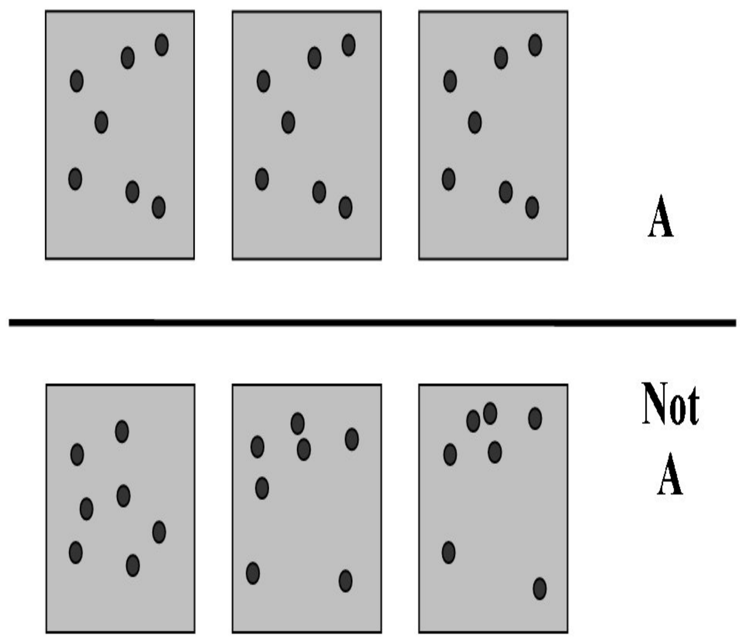

Figure 5. An example of stimuli for an (A, not A) prototype distortion task (modified from Figure 2, Ashby & Maddox, 2005).

Complex rule-based tasks can also be viewed as a form of analogy-making, as non-identical patterns need to be equated (for categorical membership) based upon some potentially abstract criteria. For example, in the Bongard problems (Bongard, 1970), a rule needs to be formed that segregates one set of figures from another (see Figure 4). Such rule generation requires selective attention (to attend to the important features) and an explicit search through hypothesis space. Thus, rule-based categorization can often require the use of several memory systems including working memory (to hold the currently active rule), semantic memory (drawing on important facts about the objects to be discriminated) and episodic memory (to hold a trace of previously attempted rules); systems central to other high-level cognitive tasks such as planning and analogy-making.

Prototype Distortion Tasks: In prototype distortion tasks, sets of stimuli are generated by first creating a prototype for each category and then generating exemplars through random distortion (e.g., Posner & Keele, 1968, 1970). To assess category learning, the categories are new to the subjects, with patterns of dots commonly being used as stimuli (e.g., D., Rhoads, & Chambliss, 1979; Homa, Sterling, & Trepel, 1981; Shin & Nosofsky, 1992; J. D. Smith & Minda, 2002). Two basic paradigms are frequently used: the (A, not A) paradigm in which subjects are given patterns belonging to category A, and are asked to decide whether or not new stimuli belong to this category (see figure 5 above); and the (A, B) paradigm in which labeled exemplars of each category are presented, with subjects requiring to categories novel items accordingly.

Both neurological and neuroimaging studies provide evidence that the (A, not A) and (A, B) paradigms rely on different neurological categorization pathways. Although in neuroimaging studies both paradigms show learning related changes in occipital cortex (indicating perceptual learning) (Reber, Stark, & Squire, 1998a; Seger et al., 2000), the (A,B) paradigm also exhibits learning related changes in activation in prefrontal and parietal cortices (indicating the use of selective attention and working memory) (Seger et al., 2000). Such findings suggest that the (A, B) paradigm relies on explicit reasoning, whereas the (A, not A) paradigm relies on low-level perceptual learning processes. This theory is supported by neuropsychological data that shows that patients with Alzhiemer's disease and amnesics show normal performance on (A, not A) tasks, but are degraded in performance on (A, B) tasks (Sinha, 1999; Knowlton & Squire, 1993; Zaki, Nosofsky, Jessup, & Unversagt, 2003).

Information-Integration Tasks: In information integration tasks, optimal categorization performance requires the integration between a number of variables (such as using a linear combination of variables to define a decision boundary). Unlike rule-based tasks, in information-integration tasks, the optimal strategy form combining variables is not easy to verbalize (such as describing the visual features that differentiate one person's face from all others).

Apart from general perceptual learning tasks (such as object recognition), information-integration tasks can involve the use of explicit variables (such as the horizontal and vertical length of line segments within a figure). In such explicit cases, it has been demonstrated that for continuous dimensions, subjects are more acute at discriminating linearly-separable classes than non-linearly separable (Ashby & Gott, 1988; Ashby & Maddox, 1990). Although there is evidence that working memory or explicit declarative memory is not involved in information-integration tasks (e.g., with Temporal Lobe Amnesics showing normal performance Filoteo, Maddox, & Davis, 2001a), striatal damage (found in Parkinson's disease) has shown to affect performance on complex (nonlinearly separable) but not simple tasks (Filoteo et al., 2001a; Ashby et al., 2003). Thus once again, depending on the nature of the task, different cognitive systems may be employed.

Weather Prediction Tasks: In weather-prediction tasks, categorical membership is probabilistic rather than deterministic. That is, during training, an example may be described as belonging to category A 60% of the time, and category B for the remaining 40%. Commonly, the experiment is conducted using tarot cards containing unique geometric patterns, with the subjects being required to judge whether the particular set of four or so cards signifies ”rain” or ”sun.” In such experiments, the rule is probabilistic, with 76% accuracy being possible. Weather prediction tasks differ from the other tasks mentioned above primarily in the variance of the strategies that subjects can use to solve the task. According to Gluck, Shohamy, and Myers (2002), subjects can utilize non-verbal information-integration rules, explicit rules, or explicit memorization. Thus, the neural categorizations pathways that are employed (and hence the models that would predict performance) vary with the individual subject, rather than being related to the task itself.

5.2 Models of Category Learning

According to Ashby and Maddox (2005), models of categorization can be divided into three main classes. Firstly, prototype models assume that a single prototype is learned for each category, with new stimuli being classified as the category with the closest matching prototype. Secondly, exemplar models assume that categories are formed by storing a set of known exemplars. In the classification of new stimuli, the similarity to each exemplar is calculated, with the stimuli being classified as the category with the highest overall similarity Nosofsky (1986). Thirdly, decision bound theories assume that the subjects partition the space into response regions. New stimuli are categorized according to which decision boundary it falls within (Maddox & Ashby, 1993). The following sections describe each of the main model types in more detail, citing exemplars, and describing which of the previously described tasks they can be used to account for.

Prototype Models: Generally speaking, prototype theories assume that for each category a single prototype is formed. The prototype is thought to represents the central tendency of a category, such as the arithmetic mean (Posner, 1969) or the mode (Neumann, 1977) of each of the features. During categorization, unseen stimuli are classified as belonging to the class with the closest matching prototype (see E. E. Smith & Medin, 1981; Medin & Smith, 1984; Homa, 1984, , for a review).

Prototype theories are best at explaining human performance at prototype distortion tasks. In such tasks it has been demonstrated that humans maximally respond to the prototypical case, even if the prototype was never studied (D. et al., 1979; Homa et al., 1981; Posner & Keele, 1968, 1970; Strange, Keeney, Kessel, & Jenkins, 1970). In such models, as only a single prototype is learned for categories A and B, a linear decision boundary will naturally form between them with stimuli falling either side being ”closer” to the segregating concept prototype. Thus, prototype theories are ill-equipped to model rule-based tasks and information-integration tasks in which the categories are non-linearly separable based on the input features.

Although standard machine learning techniques such as logistic regression, Bayesian classifiers (e.g., Duda & Hart, 1973) and 2-layered neural networks can store a single ”feature vector” for each class, generally these features do not correspond to the prototype, but to the features that discriminate between the learned categories (i.e. non-distinguishing features are often ignored). For this reason, these techniques can be viewed as decision bound models, and will be discussed later. By contrast, pure prototype models have not been utilized widely in machine learning research. Examples include forming prototypes by using typicality measures (Zhang, 1992), using Monte Carlo sampling and random mutation hill climbing (Shalak, 1994) to form a prototype, and using a top-down splitting mechanism to partition the input space into learnable prototypes (Datta & Kibler, 1995).

Exemplar Models: In exemplar models, each training stimuli and their associated category are explicitly remembered. During categorization, novel stimuli are compared to all known exemplars. Category membership can either be obtained from the best matching exemplar, the prominent response in a collection of K exemplars, or can be assigned to the category for which the sum of the similarities is greatest

Exemplar models have been proposed as a general learning mechanism, irrespective of task, and have been used to explain human performance at prototype distortion tasks, information-integration tasks and rule-based tasks. However, such a ”unified theory” does not explain the subtleties within each experimental paradigm. For example, they do not predict that maximal performance will be given to the unseen prototype in prototype distortion tasks. They also predict perfect performance on previously seen stimuli in information-integration tasks, conflicting with the experimental data. They have been used to model rule-based tasks with limited success (i.e. they cannot account for complex decisions such as those required for the Bongard problems), but assume that attention to each variable will be modified over time which will affect the similarity computations (Kruschke, 1992; Nosofsky, 1991; Nosofsky, Clark, & Shin, 1989).

With respect to exemplar models of category learning, two of the most prominent models include Nosofsky’s generalized context model (Nosofsky, 1984, 1986, 1992; Nosofsky & Zaki, 2002) and Kruschke’s Alcove model (1992). In the generalized context model, the similarity between the current example and previously stored exemplars is defined by an exponentially decreasing function based on Euclidean distance (with each feature dimension being potentially weighted). The similarities are then transformed into a choice probability, being equal to the sum of the similarities for a given class, over the sum of the similarities for all exemplars. The Alcove model extends the generalized context model by adding an error-driven learning mechanism for deriving the attentional weights assigned to each input feature dimension. Dimensional attention learning allows ALCOVE to fit human performance in situations when some stimulus dimensions are irrelevant (Shepard, Hovland, & Jenkins, 1961) or when dimensions are correlated (Medin, Altom, Edelson, & Freko, 1982).

Decision Bound Models: Decision bound theory assumes that subject partition the input space into regions that respond to each category. In labelling novel stimuli, the category can be determined from which region the stimuli falls within (Ashby & Gott, 1988; Ashby & Townsend, 1986; Maddox & Ashby, 1993).

Decision bound theories can be used to explain both prototype-distortion tasks as well as information-integration tasks. For example, Ashby and Waldron (1999) experimentally demonstrated that nonparametric models (either decision-bound or exemplar models) could be used to explain empirical studies of category leaning in information-integration tasks.

Examples of decision bound models include Ashby and Waldron’s (1999) striatal pattern classifier, Anderson (1991) rational model and Love, Medin, and Gureckis (2004) SUSTAIN (Supervised and Unsupervised Stratified Adaptive Incremental Network) model. Decision bound models are also common within general machine learning and statistical approaches to categorization. Popular approaches include the use of neural networks (for a review see Palmeri & Noelle, 2002), discriminant analysis and logistic regression (see Agresti, 1996), and support vector machines (see Christianini & Shawe-Taylor, 2000).

Hybrid Models: Although many models (such as standard machine learning techniques) exist for modelling human performance on prototype distortion and information-integration tasks, modelling the decision-making processes associated with rule-based tasks are far more complex. The most prominent model of human categorization that includes rule-based learning is the COVIS model (Competition between verbal and implicit systems) (Ashby et al., 1998). In this model it is assumed that there are two competing pathways to categorization: a non-verbal implicit pathway from the inferotemporal cortex (where the visual stimulus is represented at a high-level) through the striatum (where the stimulus is mapped to a response); and a verbal system including the anterior cingulate (where potential rules are selected), the prefrontal cortex (where the rules are evaluated), through to the striatum (where the rule parameters are learned). In this model, the pathway that provides the best performance on the given task takes control in the striatum.

Although COVIS provides an explanation of the different systems involved in different categorization tasks, mechanisms to account for complex rule-based categorization (e.g., the Bongard problems) have not been explored. As stated, such a system would need to integrate several memory systems (procedural memory, working memory, semantic memory and episodic memory), utilizing the same processes involved in other high-level perceptual tasks such as planning and analogy-making. Prominent models of such high-level problem solving include SOAR (Laird, Rosenbloom, & Newell, 1986; Laird, Newell, & Rosenbloom, 1987; Newell, 1990), ACT-R (Anderson, 1983, 1993; Anderson & Lebiere, 1998), models by the Fluid Analogies Research Group (Hofstadter, 1995), and DUAL (Kokinov, 1994).

6. Future Directions

We believe that the area of human operator modeling may be about to undergo a paradigm shift. Current academic and commercial systems are based to a large degree on process models akin to ACT (Anderson & Lebiere, 1998) and SOAR (Newell, 1990). These systems have proven brittle and require constant attention in order to capture even fairly minor variations in the tasks they are attempting to describe (Dennis, 2005).

An alternative is to treat operator modeling as a machine learning task (Quesada, Kintsch, & Gomez, 2003b, 2003a). This approach emphasizes the collection of large datasets. Rather than attempting to hand craft the representations and control processes that constitute the model, the objective is to infer these from data using the impressive arsenal of techniques available in the statistical learning literature (c.f. Hastie, Tibshirani, & Friedman, 2001). Quesada has demonstrated that at least in the case of representation learning adopting this approach has significant advantages. It has been shown to be capable of predicting the grades that flight instructors give to landings and to account for performance in simulator exercises of firefighting and plant operation (Quesada et al., 2003b, 2003a). One particularly useful property of the approach is that by varying the amount of data given to the inference techniques one can capture differences between novice and expert operators.

Within this general approach, there remain a number of areas which will require additional attention in order to make the current memory modeling literature more applicable to human operator simulation. In particular, we will comment on the issues of content, context and control. How can we build models that employ meaningful representations of stimuli (Steyvers et al., 2004; Kwantes, in press)? What is context, how can we characterize it computationally? How can we build systematic models of the control of memory rather than the current state of the art where control mechanisms are constructed on a task by task basis (Dennis, 2005)?

6.1 Content

Memory models have traditionally used arbitrary and/or hand crafted representations of the material to be learned. Thus words might be represented as vectors with random elements drawn from a particular distribution. If it were necessary to represent words which were assumed to be similar (e.g., members of the same taxonomic category) then the vectors might be crafted to contain a common set of features plus a random component. As outlined above, the same has generally been true of systems for operator modelling.

With the development of connectionism it became clear that it is possible to infer representations from data (Elman, 1991). However, typically these systems have proven difficult to apply to large representative datasets as a consequence of their computational complexity. As outlined in the section on lexical semantics above, what has been more successful is the use of machine learning techniques such as the singular value decomposition). Both the work on Latent Semantic Analysis (Landauer & Dumais, 1997) and Latent Problem Solving Analysis (Quesada et al., 2003b, 2003a) has shown that it is possible to induce representations of real world tasks that can be used to predict behavior on a fine scale.

In addition, this approach has been imported back into the memory literature. Steyvers et al. (2004) showed that representations constructed by taking the singular value decomposition of free association norms were able to account for the effects of semantic similarity on episodic tasks, such as recognition, cued recall and free recall. Furthermore, (Kwantes, in press) has shown how representations similar to those produced by machine learning techniques can be created using memory mechanisms that are more psychologically plausible. This work is important as it strengthens the rational behind the use of the machine learning techniques.

Sufficient progress has been made in the induction of content to provide confidence that approaches based on this approach will be viable. However, much remains to be achieved. In particular, current work has focused on constructing the meanings of individual words or actions. Clearly, however, systems capable of capturing relational information will ultimately be an important component of operator modeling systems. The work on the Syntagmatic Paradigmatic model (Dennis, 2005) is an important first step in understanding how such information could be induced, but more work is necessary.

6.2 Context

A limitation in applying the episodic memory models to applied situations is the difficulty in specifying episodes and events within episodes in the real world. Many contemporary models of human memory utilize fairly simplistic representational methods that presume a simple drift of the contextual vector (Mensink & Raaijmakers, 1988; Dennis & Humphreys, 2001), although see (Howard & Kahana, 2002) for an exception. In unified models, like ACT-R, an episode is often represented by a separate node in a spreading activation network. However, episodes are probably defined by the goals and intentions of the person. Because there will typically be more than one goal active at any time and because events may be related to several goals any simple partitioning of the event stream or constant drift mechanism is unlikely to be a sufficient characterization of contextual change. For example, when watching a movie, there are no clear cut ”episodes”. Rather there are plots and subplots each of which overlap and interleave to a significant degree.

Horvitz, Dumais, and Koch (2004) provide one line of work which provides some insight into how one might approach the problem of contextual change within a machine learning framework. They focused on identifying memory landmarks in an email stream using a Bayesian network and a support vector machine to build classifiers based on supervised training data. A similar method could be used to identify the starts and ends of episodes in operator action streams. Further work is necessary, however, to determine the usefulness of such a system.

6.3 Control

The area of control is one of the most difficult questions in human operator modeling and in cognitive science as a whole. The limitation of the current generation of models in which the control over the task the model performs is maintained by the modeler not the model’s history, or the current environment in which the model is operating is a serious impediment to progress.

One reason that progress has been slow in this area is the difficulty in finding empirical paradigms that shed light on the nature of the control mechanisms at play in any given task. Humphreys (personal communication) has suggested that one possible approach would be to adapt task switching paradigms to look for priming between tasks. To the extent that a task is influenced by preceding tasks across changes in materials etc. we could infer that elements of the control architecture of the first task continued to be active during the execution of the second task. This seems like a promising approach to a difficult issue.

A second reason that progress has been slow is that it has been unclear how the issue of control might be addressed theoretically. In an important first step, Dennis (2005) showed in the context of verbal semantic memory how instructions can be incorporated as an input to the model to capture how subjects may switch between rating and categorization tasks and how they can use the same background information in support of each. The important development in this work is that the two separate tasks are not modeled independently but arise as emergent behaviors from a single memory based system. However, the tasks employed are simple and more comprehensive demonstrations will be necessary.

7. Conclusions

We suspect that the area of human operator modelling is about to undergo a transformation. Rather than the large finely crafted models that predominate today, we will see a generation of far simpler models that apply machine learning techniques to large data sets of operator actions. To increase the fidelity of these models, it will continue to be important to consider key memory phenomena and models in the areas of short term memory, semantic memory, episodic memory, prospective memory and categorization. Furthermore, we believe that human operator modeling provides a useful context to guide future empirical work and the next generation of memory models. In particular, working in this context of use emphasizes the need to address fundamental issues of content, context and control.

8. Author Notes

Please address correspondence to Dr. Simon Dennis, Department of Psychology, University of Adelaide. Email: Simon.Dennis@adelaide.edu.au.

Footnotes 1Although, in recent work, (Quesada et al., 2003b, 2003a) have applied language based models such as Latent Semantic Analysis to series of operator actions as provided in simulation contexts. We believe this work has the potential to transform the way in which operator models are constructed and deserves close attention.

9. References

Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A learning algorithm for boltzmann machines. Cognitive Science, 9, 147-169.

Agresti, A. (1996). An introduction to categorical data analysis. New York: John Wiley.

Anderson, J. R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press. Anderson, J. R. (1991). The adaptive nature of human categorization. Psychological Review, 98, 409-29.

Anderson, J. R. (1993). Rules of the mind. Hillsdale, NJ: Erlbaum.

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. Journal of Memory and Language, 38, 341-380.

Anderson, J. R., & Bower, G. H. (1973). Human associative memory. Oxford, England: V. H. Winston & Sons.

Anderson, J. R., & Lebiere, C. (1998). The atomic components of thought. Mahwah, New Jersey: Lawrence Erlbaum associates.

Anderson, J. R., & Matessa, M. (1997). A production system theory of serial memory. Psychological Review, 104, 728-748.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldon, E. M. (1998). A neuropsycholoigcal theory of multiple systems in category learning. Psychological Review, 105, 442-481.

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. Journal of Experimental Psychology: Learning, Memory and Cognition, 14, 33-53.

Ashby, F. G., & Maddox, W. T. (1990). Integrating information from separable psychological dimensions. Journal of Experimental Psychology: Human Perception and performance, 16, 598-612.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. Annual Review of Psychology, 56, 149-178.

Ashby, F. G., Noble, S., Filoteo, J., Waldron, E. M., & Ell, S. W. (2003). Category learning deficits in parkinson’s disease. Neuropsychology, 17, 115-24.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. Psychological Review, 93, 154-79.

Ashby, F. G., & Waldron, E. M. (1999). On the nature of implicit categorization. Psychonomic Bulletin and Review, 6, 363-78.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), Recent advances in learning and motivation (volume 8). New York: Academic Press.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. Journal of Verbal Learning and Verbal Behavior, 14, 575-589.

Baddeley, A. D., & Wilson, B. (1984). Phonological coding and short-term memory in patients without speech. Journal of Memory and Language, 24, 490-502.

Blaheta, D., & Charniak, E. (2000, May). Assigning function tags to parsed text. In Proceedings of the 1st annual meeting the north american chapter of the acl (naacl) (p. 234-240). Seattle, Washington.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2002). Latent dirichlet allocation. In Nueral information porcessing systems (Vol. 14). Lawrence Erlbaum Associates.

Bongard, M. (1970). Pattern recognition. New York: Spartan Books.

Bower, G. H. (2000). A brief history of memory research. In E. Tulving & F. I. M. Craik (Eds.), The oxford handbook of memory (p. 3-32). Oxford: Oxford University Press.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. Psychological Review, 107, 127181.

Brown, R. G., & Marsden, C. D. (1988). Internal versus external cues and the control of attention in parkinson's disease. Brain, 111, 323-45.

Burgess, P. W., & Shallice, T. (1997). The relationship between prospective memory and retrospective memory: Neuropsychological evidence. In M. A. Conway (Ed.), Cognitive models of memory (p. 247-272). Cambridge, MA: MIT.

Chappell, M., & Humphreys, M. S. (1994). An auto-associative neural network for sparse representations: Analysis and application to models of recognition and cued recall. Psychological Review, 101, 103-128.

Christianini, N., & Shawe-Taylor, J. (2000). An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press. Collins, A., &

Loftus, E. (1975). A spreading activation theory of semantic memory. Psychological Review, 82, 407-428.

Collins, M., & Quillian, M. R. (1969). Retrieval from semantic memory. Journal of Verbal Learning and Verbal Behavior, 8, 240-247.

Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. Cognitive Science, 23, 371-414.

Criss, A. H., & Shiffrin, R. M. (2004). Context noise and item noise jointly determine recognition memory: a comment on dennis and humphreys (2001). Psychological Review, 111, 800-807.

D., Rhoads, D., & Chambliss, D. (1979). Evolution of conceptual structure. Journal of Experimental Psychology: Learning, Memory and Cognition, 5, 11-23.

Datta, P., & Kibler, D. (1995). Learning prototypical concept descriptions. In Machine learning: Proceedings of the 12 th international conference icml-95. San Francisco, CA, USA.

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., & Haarmann. (2005). The demise of short-term memory revisited: Empirical and computational investigations of recency effect. Psychological Review, 112, 3-42.

Dennis, S. (2004). An unsupervised method for the extraction of propositional information from text. Proceedings of the National Academy of Sciences, 101, 5206-5213.

Dennis, S. (2005). A memory-based theory of verbal cognition. Cognitive Science, 29 (2), 145-193.

Dennis, S., & Humphreys, M. S. (2001). A context noise model of episodic word recognition. Psychological Review, 108 (2), 452-478.

Dennis, S., Jurafsky, D., & Cer, D. (2003). Supervised and unsupervised models for propositional analysis. In Workshop on syntax, semantics and statistics at the neural information processing society conference. Vancouver, BC.

Duda, R. O., & Hart, P. E. (1973). Pattern classification and scene analysis. New York: Wiley.

Eich, J. M. (1982). A composite holographic associative recall model. Psychological Review, 89 (6), 627-661.

Einstein, G. O., & McDaniel, M. A. (1990). Normal aging and prospective memory. Journal of Experimental Psychology:Learning, Memory, and Cognition, 16, 717-726.

Einstein, G. O., & McDaniel, M. A. (1996). Retrieval processes in prospective memory: Theoretical approaches and some new empirical findings. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), Prospective memory: Theory and application (p. 115-141). Hillsdale, NJ: Lawrence Erlbaum Associates.

Einstein, G. O., McDaniel, M. A., Manzi, M., Cochran, B., & Baker, M. (2000). Prospective memory and aging: Forgetting intentions over short delays. Psychology and Aging, 15, 671-683.

Einstein, G. O., McDaniel, M. A., Williford, C. L., Pagan, J. L., & Dismukes, R. K. (2003). Forgetting of intentions in demanding situations in rapid. Journal of Experimental Psychology: Applied, 3, 147-162.

Ellis, J. A. (1996). Prospective memory or the realization of delayed intentions: A conceptual framework for research. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), Prospective memory: Theory and application (p. 1-22). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ellis, J. A., & Milne. (1996). Retrieval cue specificity and the realization of delayed intentions. Journal of Experimental Psychology: Human Experimental Psychology, 49(A), 862-877.

Elman, J. L. (1991). Distributed representations, simple recurrent networks and grammatical structure. Machine Learning, 7, 195-225.

Estes, W. K. (1991). On types of item coding and source of recall in short-term memory. In W. E. Hockley & S. Lewandosky (Eds.), Relating theory and data: Essays on human memory in honor of Bennet B. Murdock (p. 155-174). Hillsdale, NJ: Erlbaum.

Fillmore, C. J. (1971). Some problems for case grammar. In R. J. O'Brien (Ed.), 22nd round table. linguistics: developments of the sixties - viewpoints of the seventies (Vol. 24, p. 35-56). Washington D. C.: Georgetown University Press.

Fillmore, C. J., Wooters, C., & Baker, C. F. (2001). Building a large lexical databank which provides deep semantics. In Proceedings of the pacific asian conference on language, information and computation. Hong Kong.

Filoteo, J. V., Maddox, W. T., & Davis, J. (2001a). A possible role of the striatum in linear and nonlinear categorization rule learning: evidence from patients with huntington's disease. Behavioural Neuroscience, 115, 786-98.

Flexser, A. J., & Tulving, E. (1978). Retrieval independence in recognition and recall. Psychological Review, 85 (3), 153-171.

Frege, G. (1879). Begriffsschrift, eine der arthimetischen nachgebildete formelsprache des reinen denkens. Halle: Nebert. (Translated by T. W. Bynum, Conceptual notation and related articles. Oxford: Claredon Press, 1972.)

Ge, X., & Iwata, S. (2002). Learning the parts of objects by auto-association. Neural Networks, 15, 285-295.

Gildea, D., & Jurafsky, D. (2002). Automatic labeling of semantic roles. Computational Linguistics, 28 (3), 245-288.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. Psychological Review, 91 (1), 1-67.

Gionis, A., Indyk, P., & Motwani, R. (1999). Similarity search in high dimensions via hashing. In The VLDB journal (p. 518-529).

Gluck, M. A., Shohamy, D., & Myers, C. (2002). How do people solve the ”weather prediction” task? individual variability in strategies for probabilistic category learning. Learning and Memory, 9, 408-18.

Goldman-Rakic, P. S. (1987). Circuitry of the prefrontal cortex and the regulation of behaviour by representational knowledge. In F. Plum & V. Mountcastle (Eds.), Handbook of physiology (p. 373-417). Bethesda, MD: American Physiological Society. Goldman-Rakic, P. S. (1995). Cellular basis of working memory. Neuron, 14, 477-85.

Griffths, T. L., & Steyvers, M. (2002). Prediction and semantic association. In Nips.

Guynn, M. J. (2003). A two-process model of strategic monitoring in event-based prospective memory: Activation/retrieval mode and checking. International Journal of Psychology, 38, 245-256.

Guynn, M. J., McDaniel, M. A., & Einstein, G. O. (2001). Remembering to perform actions: A different type of memory? In H. D. Zimmer, R. L. Cohen, M. J. Guynn, J. Engelkamp, R. Kormi-Nouri, & M. A. Foley (Eds.), Memory for action: A distinct form of episodic memory? (p. 25-48). New York: Oxford University Press.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). The elements of statistical learning: Data mining, inference and prediction. New York: Springer.

Haykin, S. (1994). Neural networks. New York: Macmillan College Publishing Company, Inc.

Healy, A. F., & McNamara, D. S. (1996). Verbal learning and memory: Does the modal model still work? Annual Review of Psychology, 47, 143-172.

Heaton, R. K. (1981). A manual for the wisconsin card sorting test. Odessa, FL: Psychol. Assess. Resourc.

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. E. Delafresnaye (Ed.), Brain mechanisms and learning (p. 37-46). New York: Oxford University Press.

Henson, R. N. A. (1998). Short-term memory for serial order: The start-end model. Cognitive Psychology, 36, 73-137.

Hintzman, D. L. (1984). Minerva-2 - a simulation-model of human-memory. Behavior Research Methods Instruments & Computers, 16 (2), 96-101.

Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal grouping effects in immediate recall: A working memory analysis. Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 49A, 116-139.

Hockley, W. E. (1992). Item versus associative information : Further comparisons of forgetting rates. Journal of Experimental Psychology: Learning, Memory and Cognition, 18 (6), 1321-1330.

Hofmann, T. (2001). Unsupervised learning by probabilistic latent semantic analysis. Machine Learning, 42 (1-2), 177-196.

Hofstadter, D. (1995). Fluid concepts and creative analogies: Computer models of fundamental mechanisms of thought. New York: Basic Books.

Homa, D. (1984). On the nature of categories. In G. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory (p. 49-94). San Diego, CA: Academic Press.

Homa, D., Sterling, S., & Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. Journal of Experimental Psychology: Learning, Memory and Cognition, 7, 418-39.

Hopfield, J. J. (1982). Neural networks as physical systems with emergent computational abilities. Proceedings of the National Academy of Sciences of the USA, 79, 2554-58.

Horvitz, E., Dumais, S., & Koch, P. (2004). Learning predictive models of memory landmarks. In Proceedings of the 26th annual meeting of the cognitive science society. Chicago, IL.

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. Journal of Mathematical Psychology, 46, 268-299.

Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system - a theory for episodic, semantic, and procedural tasks. Psychological Review, 96 (2), 208-233.

Humphreys, M. S., Dennis, S., Maguire, A. M., Reynolds, K., Bolland, S. W., & Hughes, J. D. (2003). What you get out of memory depends of the question you ask. Journal of Experimental Psychology: Learning, Memory and Cognition, 29 (5), 797-812.

Humphreys, M. S., Pike, R., Bain, J. D., & Tehan, G. (1989). Global matching - a comparison of the sam, minerva-ii, matrix, and todam models. Journal of Mathematical Psychology, 33 (1), 36-67.

Humphreys, M. S., Tehan, G., O'Shea, A., & Boland, S. W. (2000). Target similarity effects: Support for the parallel distributed processing assumptions. Memory and Cognition, 28, 798-811.

Humphreys, M. S., Weeks, C. S., & Hockley, W. E. (n.d.). Buffered forgetting: What happens when target and distracter matching strength both decline?

Humphreys, M. S., Wiles, J., & Dennis, S. (1994). Toward a theory of human-memory - data-structures and access processes. Behavioral and Brain Sciences, 17 (4), 655-667.

Isbell, C. L., & Viola, P. (1999). Restructuring sparse high dimensional data for effective retrieval. In Advances in neural information processing systems conference (Vol. 11, p. 480-486). MIT Press, Cambridge MA.

Jackendoff, R. S. (1992). Languages of the mind: Essays on mental representation. Cambridge, MA: MIT Press.

Kahana, M. J. (1996). Associative retrieval processes in free recall. Memory and Cognition, 24, 103109.

Keppel, G., & Underwood, B. (1962). Proactive inhibition in short-term retention of single items. Journal of Verbal Learning and Verbal Behavior, 1, 153-161.

Kimberg, D. Y., D'Esposito, M., & Farah, M. J. (1997). Frontal lobes: neuropsychological aspects. In T. Feinberg (Ed.), Behavioral neurology and neuropsychology (p. 409-18). New York: McGraw-Hill.

Kingsbury, P., Palmer, M., & Marcus, M. (2002). Adding semantic annotation to the penn treebank. In Proceedings of the human language technology conference. San Diego, CA.

Kintsch, W. (1998). Comprehension: A paradigm for cognition. Cambridge University Press.

Kliegal, M., Martin, M., McDaniel, M. A., & Einstein, G. O. (2004). Importance effects on performance in event-based prospective memory tasks. Memory, 12, 553-561.

Knowlton, B. J., & Squire, L. R. (1993). The learning of natural categories: parallel memory systems for item memory and category-level knowledge. Science, 262, 1747-49.

Kokinov, B. (1994). The dual cognitive architecture: A hybrid multi-agent approach. In A. Cohn (Ed.), Proceedings of the eleventh european conference on artificial intelligence. London: John Wiley & Sons

Konishi, S., Karwazu, M., Uchida, I., Kikyo, H., Asakura, I., & Miyashita, Y. (1999). Contribution of working memory to transient activation in human inferior prefrontal cortex during performance ofthe wisconsinn cardsorting test. Cerebral Cortex, 9, 745-73.

Kruschke, J. K. (1992). Alcove: An exemplar-based connectionist model of category learning. Psychological Review, 99, 22-44.

Kwantes, P. J. (in press). Using context to build semantics. *Psychonomic Bulletin and Review.*

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). Soar: An architecture for general intelligence. Artificial Intelligence, 33, 1-64.

Laird, J. E., Rosenbloom, P. S., & Newell, A. (1986). Chunking in soar: The anatomy of a general learning mechanism. Machine Learning, 1, 11-46.

Lakoff, G. (1987). Women, fire and dangerous things: What categories reveal about the mind. Chicago: University of Chicago Press.

Landauer, T. K., & Dumais, S. T. (1997). A solution to plato's problem: The latent semantic analysis theory of the acquisition, induction, and representation of knowledge. Psychological Review, 104, 211-240.

Langacker, R. W. (1987). Foundations of cognitive grammar 1: Theoretical prerequisites. Stanford, CA: Stanford University Press.

Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. Nature, 401 (6755), 788-791.

Lewandowsky, S., & Murdock, B. B. (1989). Memory for serial order. Psychological Review, 96 (1), 25-57.

Loft, S., Humphreys, M., & Neal, A. (2003). Prospective memory in air traffc control. In G. Edkins & P. Pfister (Eds.), Innovation and consolidation in aviation. Aldershot, UK: Ashgate.

Lombardi, W. J., Andreason, P. J., Sirocco, K. Y., Rio, D. E., & Gross, R. E. (1999). Wisconsin card sorting test performance following head injury: Dorsolateral fronto-striatal circuit activity predicts perseveration. Journal of Clinical and Experimental Neuropsychology, 21, 2-16.

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). Sustain: a network model of category learning. Psychological Review, 111, 309-32.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co- occurrence. Behavior Research Methods Instruments & Computers, 28 (2), 203-208.

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. Perception and Psychophysics, 53, 49-70.

Maguire, A. M., Humphreys, M. S., & Dennis, S. (in preparation). False alarms in episodic recognition: An examination of base-rate, similarity-based, and comprehensive theories.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. Psychological Review, 87 (3), 252-271.

Marsh, R. L., & Hicks, J. L. (1998). Event-based prospective memory and executive control of working memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 24, 336-349.

Marsh, R. L., Hicks, J. L., & Cook, G. I. (2005). On the relationship between effort toward an ongoing task and cue detection in event-based prospective memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 31, 68-75.

Marsh, R. L., Hicks, J. L., Cook, G. I., Hansen, J. S., & Pallos, A. L. (2003). Interference to ongoing activities covaries with the characteristics of an event-based intention. Journal of Experimental Psychology: Learning, Memory, and Cognition, 29, 861-870.

Marsh, R. L., Hicks, J. L., & Hancock, T. W. (2000). On the interaction of ongoing cognitive activity and the nature of an event-based intention. Applied Cognitive Psychology, 14, S29-S41.

Marzal, A., & Vidal, E. (1993). Computation of normalized edit distance and applications. IEEE trans. on Pattern Analysis and Machine Intelligence, 15 (9), 926–932.

McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. Psychological Review, 105 (4), 724-760.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. Psychological Review, 102 (3), 419-457.

McDaniel, M. A. (1995). Prospective memory: Progress and process. Psychology of Learning and Motivation, 33, 191-221.

McDaniel, M. A., & Einstein, G. O. (1993). The importance of cue familiarity and cue distinctiveness in prospective memory. Memory, 1, 23-41.

McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. Applied Cognitive Psychology, 14, 127-144.

McDaniel, M. A., Einstein, G. O., Graham, T., & Rall, E. (in press). Delaying execution of intentions. overcoming the costs of interruptions. Applied Cognitive Psychology, 18 (5), 553-547.

McDaniel, M. A., Einstein, G. O., Stout, A. C., & Morgan, Z. (2003). Aging and maintaining intentions over delays: Do it or lose it. Psychology and Ageing, 18 (4), 823-835.

McDaniel, M. A., Einstein, M. J. G. G. O., & Breneiser, J. (2004). Cue-focussed and automatic-associative processes in prospective memory retrieval. Journal of Experimental Psychology: Learning, Memory, and Cognition, 30, 605-614.

Medin, D. L., Altom, M. W., Edelson, S. M., & Freko, D. (1982). Correlated symptoms and simulated medical classification. Journal of Experimental Psychology: Learning, Memory and Cognition, 8, 37-50.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. Annual Review of Psychology, 35, 113-138. Meier, B., & Graf, P. (2000). Transfer appropriate processing for prospective memory tests. Applied Cognitive Psychology, 14, 11-27.

Mensink, G. J., & Raaijmakers, J. G. W. (1988). A model for interference and forgetting. Psychological Review, 95 (4), 434-455.

Moscovitch, M. (1992). Memory and working with memory: A component process model based on modules and central systems. Journal of Cognitive Neuroscience, 4, 257-267.

Murdock, B. B. (1968). Serial order effect in short term memory. Journal of Experimental Psychology, 76, 1-15.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. Psychological Review, 89 (6), 609-626.

Murdock, B. B. (1983). A distributed memory model for serial-order information. Psychological Review, 90 (4), 316-338.

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. Annual Review of Psychology, 53, 53-81.

Neath, I., & Surprenant, A. M. (2003). Human memory: An introduction to research, data and theory (2 ed.). Wadsworth.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), Basic processes in reading (p. 264336). Hillsdale, NJ: Lawrence Erlbaum Associates.

Neumann, P. G. (1977). Visual prototype formation with discontinuous representation of dimensions of variability. Memory and Cognition, 5, 187-197.

Newell, A. (1990). The unified theories of cognition. Cambridge, MA: Harvard University Press.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. Journal of Experimental Psychology: Learning, Memory and Cognition, 10, 104-114.
Nosofsky, R. M. (1986). Attention, similarity and the idenitification-categorization relationship. Journal of Experimental psychology: General, 115, 39-57.

Nosofsky, R. M. (1991). Typicality in logically defined categories: exemplar-similarity versus rule instantiation. Memory and Cognition, 19, 131-50.

Nosofsky, R. M. (1992). Exemplars, prototypes and similarity rules. In A. F. Healy, S. M. Kosslyn, & R. M. Shiffrin (Eds.), From learning theory to connectionist theory: Essays in honour of william k. estes vol. 1. Hillsdale, NJ: Lawrence Erlbaum.

Nosofsky, R. M., Clark, S. E., & Shin, H. J. (1989). Rules and exemplars in categorization, identification, and recognition. Journal of Experimental Psychology: Learning, Memory and Cognition, 15, 282-304.

Nosofsky, R. M., & Zaki, S. R. (2002). Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. Journal of Experimental Psychology: Learning, Memory and Cognition, 28, 924-940.

O'Hara, T., & Wiebe, J. (2002). Classifying preposition semantic roles using class-based lexical associations (Tech. Rep. No. NMSU-CS-2002-013). Computer Science Department, New Mexico State University.

O'Reilly, R. C., & Munakata, Y. (2000). Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain. Cambridge, MA: MIT Press.

O'Reilly, R. C., Norman, K. A., & McClelland, J. L. (1998). A hippocampal model of recognition memory. In M. I. Jordan, M. J. Kearns, & S. A. Solla (Eds.), Advances in neural information processing systems 10 (p. 73-79). Cambridge, MA: MIT Press.

Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. Psychological Review, 105 (4), 761-781.

Palmer, M., Rosenzweig, J., & Cotton, S. (2001). Automatic predicate argument analysis of the penn treebank. In J. Allan (Ed.), Proceedings of hlt 2001, first international conference on human language technology research. San Francisco: Morgan Kaufmann.

Palmeri, T. J., & Noelle, D. (2002). Concept learning. In M. A. Arbib (Ed.), The handbook of brain theory and neural networks. MIT Press.

Pew, R. W., & Mavor, A. S. (1998). Modeling human and organizational behavior: Applications to military simulations. National Academy Press.

Plaut, D. C. (1995). Semantic and associative priming in a distributed attractor network. In Proceedings of the 17th annual conference of the cognitive science society. Hillsdale, NJ: Lawrence Erlbaum Associates.

Posner, M. I. (1969). Abstraction and the process of recognition. In G. H. Bower & J. T. Spence (Eds.), The psychology of learning and motivation. vol. 3. New York, NY: Academic Press.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. Journal of Experimental Psychology, 77, 353- 63.

Posner, M. I., & Keele, S. W. (1970). Retention of abstract ideas. Journal of Experimental Psychology, 83, 304-8.

Posner, M. I., & Petersen, S. E. (1990). Attention systems in the human brain. Annual Review of Neuroscience, 13, 25-42.

Pradhan, S., Hacioglu, K., Ward, W., Martin, J., & Jurafsky, D. (2003). Shallow semantic parsing using support vector machines (Tech. Rep. No. TR-CSLR-2003-03). Center for Spoken Language Research, University of Colorado.

Quesada, J., Kintsch, W., & Gomez, E. (2003a). Automatic landing technique assessment using latent problem solving analysis. In R. Alterman & D. Kirsh (Eds.), Twenty fifth conference of the cognitive science society (Vol. 25). Boston, MA: Lawrence Erlbaum Associates.

Quesada, J., Kintsch, W., & Gomez, E. (2003b). Latent problem solving analysis as an explanation of expertise effects in a complex, dynamic task. In R. Alterman & D. Kirsh (Eds.), Twenty fifth conference of the cognitive science society (Vol. 25). Boston, MA: Lawrence Erlbaum Associates.

Raijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. Psychological Review, 88 (2), 93-134.

Raaijmakers, J. G. W., & Shiffrin, R. M. (2002). Models of memory. In H. Pashler & D. Medin (Eds.), Stevens' handbook of experimental psychology: Memory and cognitive processes (Vol. 2). New York: John Wiley Sons Inc.

Rao, S. M., Bobholz, J. A., Hammeke, T. A., Tosen, A. C., & Woodley, S. J. (1997). Functional mri evidence for subcortical participation in conceptual reasoning skills. Neuroreport, 8, 1987- 93.

Ratcliff, R., Clark, S. E., & Shiffrin, R. M. (1990). List-strength effect .1. data and discussion. Journal of Experimental Psychology-Learning Memory and Cognition, 16 (2), 163-178.

Reber, P. J., Stark, C. E. L., & Squire, L. R. (1998a). Contrasting cortical activity associated with category memory and recognition memory. Learning and Memory, 5, 420-28.
Rogers, R. D., Andrews, T. C., Grasby, P. M., Brooks, D. J., & Robbins, T. W. (2000). Contrasting cortical and subcortical activations produced by attentional-set shifting and reversal learning in humans. Journal of Cognitive Neuroscience, 12, 142-62.

Roweis, S. T., & Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. Science, 290 (5500), 2323-+.

Seger, C. A., Poldrack, R. A., Prabhakaran, V., Zhao, M., Glover, G. H., & Gabrieli, J. D. E. (2000). Hemispheric asymmetries and individual differences in visual concept learning as measured by functional mri. Neuropsychologia, 38, 1316-24.

Seung, H. S. (1998). Learning continuous attractors in recurrent networks. Advances in Neural Information Processing Systems, 11.

Shalak, D. (1994). Prototype and feature selection by sampling and random mutation hill climbing algorithms. In Proceedings of the eleventh international machine learning conference. New Brunswick, NJ: Morgan Kaufmann.

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. Psychological Monographs, 13.

Shiffrin, R. M., Ratcliff, R., & Clark, S. E. (1990). List-strength effect .2. theoretical mechanisms. Journal of Experimental Psychology-Learning Memory and Cognition, 16 (2), 179-195.

Shiffrin, R. M., & Steyvers, M. (1997). Model for recognition memory: Rem - retrieving effectively from memory. Psychonomic Bulletin & Review, 4 (2), 145-166.

Shin, H. J., & Nosofsky, R. M. (1992). Similarity scaling studies of dot-pattern classification and recognition. Journal of Experimental Psychology: General, 121, 278-304.

Sinha, R. R. (1999). Neuropsychological substrates of category learning. Unpublished doctoral dissertation.

Slamecka, N. J. (1966). Differentiation versus unlearning of verbal associations. Journal of Experimental Psychology, 71, 822 - 828.

Slonim, N., & Tishby, N. (2000). Document clustering using word clusters via the information bottleneck method. In Research and development in information retrieval (p. 208-215).

Smith, E. E., & Medin, D. L. (1981). Categories and concepts. Cambridge, MA: Harvard University Press.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A feature model for semantic decision. Psychological Review, 81, 214-241.

Smith, J. D., & Minda, J. P. (2002). Distinguishing prototype-based and exemplar-based processes in category learning. Journal of Experimental Psychology: Learning, Memory and Cognition, 28, 800-11.

Smith, R. E. (2003). The cost of remembering to remember in event-based prospective memory: Investigating the capacity demands of delayed intention performance. Journal of Experimental Psychology: Learning, Memory, and Cognition, 29, 347-361.

Smith, R. E., & Bayen, U. J. (2004). A multinomial model of prospective memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 30, 756-777.

Stallard, D. (2000). Talk'n'travel: A conversational system for air travel planning. In Proceedings of the 6th applied natural language processing conference (anlp'00) (p. 68-75).

Steyvers, M., Shiffrin, R. M., & Nelson, D. L. (2004). Word association spaces for predicting semantic similarity effects in episodic memory. In A. Healy (Ed.), Experimental cognitive psychology and its applications: Festschrift in honor of lyle bourne, walter kintsch, and thomas landauer. Washington, DC: American Psychological Association.

Strange, W., Keeney, T., Kessel, F. S., & Jenkins, J. J. (1970). Abstraction over time of prototypes from distortions of random dot patterns: a replication. Journal of Experimental Psychology, 83, 508- 10.

Tulving, E. (1976). Ecphoric processes in recall and recognition. In J. Brown (Ed.), Recall and recognition (p. 275). Oxford, England: John Wiley & Sons.

Tulving, E. (1983). Elements of episodic memory. Oxford: Clarendon Press.

Van Valin, R. D. (1993). A synopsis of role and reference grammar. In R. D. Van Valin (Ed.), Advances in role and reference grammar. Amsterdam: John Benjamins Publishing Company.

Volz, H. P., Gaser, C., Rzanny, F. H. R., & Mentzel, H.-J. (1997). Brain activation during cognitive stimulation with the wisconsin card sorting test-a functional mri study on healthy volunteers and schizophrenics. Psychiatry Research: Neuroimaging, 75, 45- 157.

West, R., Krompinger, J., & Bowry, R. (in press). Disruptions of preparatory attention contribute to failures of prospective memory. Psychonomic Bulletin and Review.

Wiles, J., Humphreys, M. S., Bain, J. D., & Dennis, S. (1991). Direct memory access using two cues: Finding the intersection of sets in a connectionist model. In R. P. Lippman, J. E. Moody, & T. D. S. (Eds.), Advances in neural information processing systems 3 (p. 635-641). San Mateo, CA: Morgan Kaufmann.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. Journal of Memory and Language, 46 (3), 441-517.

Zaki, S. R., Nosofsky, R. M., Jessup, N. M., & Unversagt, F. W. (2003). Categorization and recognition performance of a memory-impaired group: evidence for single-system models. Journal of the International Neuropsychological Society, 9, 394-406.

Zhang, J. (1992). Selecting typical instances in instance-based learning. In Proceedings of the international machine learning conference 1992 (p. 470-479).

DOCUMENT CONTROL DATA SHEET

<table>
<tr><td>1a. PERFORMING AGENCY
University of Adelaide</td><td rowspan="2">2. SECURITY CLASSIFICATION

UNCLASSIFIED
–</td></tr>
<tr><td>1b. PUBLISHING AGENCY
DRDC Toronto</td></tr>
<tr><td colspan="2">3. TITLE

(U) Human Memory Models for Operator Simluation /</td></tr>
<tr><td colspan="2">4. AUTHORS

Simon Dennis; Michael S Humphreys; Scott W Boland; Steven Savvas; Shayne D Loft</td></tr>
<tr><td>5. DATE OF PUBLICATION

June 10 , 2005</td><td>6. NO. OF PAGES

58</td></tr>
<tr><td colspan="2">7. DESCRIPTIVE NOTES

Review of the Human Memory Literature of the purpose of supporting our the HBR programme at DRDC Toronto.</td></tr>
<tr><td colspan="2">8. SPONSORING/MONITORING/CONTRACTING/TASKING AGENCY
Sponsoring Agency:
Monitoring Agency:
Contracting Agency : DRDC Toronto
Tasking Agency:</td></tr>
</table>

<table>
<tr><td>9. ORIGINATORS DOCUMENT NO.

Contract Report CR 2005−185</td><td>10. CONTRACT GRANT AND/OR PROJECT NO.

Contract #
W7711−047931/001/TOR</td><td>11. OTHER DOCUMENT NOS.</td></tr>
<tr><td colspan="3">12. DOCUMENT RELEASABILITY

Unlimited distribution</td></tr>
<tr><td colspan="3">13. DOCUMENT ANNOUNCEMENT

Unlimited announcement</td></tr>
</table>

14. ABSTRACT

(U) In this review of the human memory literature, we focus on those areas critical to human operator modeling and on the computational techniques from the memory modeling and machine learning literatures that are relevant to simulating human behavior in this area. We outline the areas of short term memory, semantic memory, episodic memory, prospective memory and categorization. In addition, we review models that span these areas and memory phenomena which have yet to attract modelling efforts but which are likely to be important in operator modelling. Finally, we outline our Recommendations as to which areas will need specific attention in order to build robust models of human operators.

(U) French language abstract to be added.

15. KEYWORDS, DESCRIPTORS or IDENTIFIERS

(U) literature review; human memory; simulated operator