Part I

Introduction

1 Concepts and theories of memory

John M. Gardiner

> **Concept**. A thought, idea; disposition, frame of mind; imagination, fancy; .... an idea of a class of objects.
> **Theory**. A scheme or system of ideas or statements held as an explanation or account of a group of facts or phenomena; a hypothesis that has been confirmed or established by observation or experiment, and is propounded or accepted as accounting for the known facts; a statement of what are known to be the general laws, principles, or causes of something known or observed.
>
> From definitions given in the *Oxford English Dictionary*

The *Oxford Handbook of Memory*, edited by Endel Tulving and Fergus Craik, was published in the year 2000. It is the first such book to be devoted to the science of memory. It is perhaps the single most authoritative and exhaustive guide as to those concepts and theories of memory that are currently regarded as being most vital. It is instructive, with that in mind, to browse the exceptionally comprehensive subject index of this handbook for the most commonly used terms. Excluding those that name phenomena, patient groups, parts of the brain, or commonly used experimental procedures, by far the most commonly used terms are *encoding* and *retrieval processes*. Terms for different kinds of memory also feature prominently, as one would expect. Among the most frequently used are *short-term* and *long-term memory*; *explicit* and *implicit memory*; *working memory*; *episodic* and *semantic memory*; *verbal, visual* and *procedural memory*. All these terms may refer, among other things, to different memory systems and *memory systems theory* itself also has a lengthy entry in the index. Other commonly used terms are more disparate. They include such terms as *attention*, *consciousness*, *learning*, *forgetting*, *priming*, *recollection* and *remembering*.

It is important to distinguish between such terms and the concepts they may refer to, not least because any such term may be used in the literature to refer to several quite different concepts. One of the most notorious examples of such usage concerns explicit memory which, as Richardson-Klavehn and Bjork (1988) pointed out, has sometimes been used to refer

to the conscious state of being aware of memory, sometimes to refer to an experimental procedure, or kind of test, and sometimes to refer to a kind of memory, or memory system. Similarly, the term episodic memory is often used to refer to a class of memory tasks (which was in fact its first usage historically) as well as to a mind/brain system. Even encoding, which might seem a more straightforward term, may refer to a memory task, or to an experimental manipulation, or to a class of hypothetical processes that are assumed to be involved in the performance of a task. This 'duality of patterning' in the usage of terminology has sometimes led to conceptual confusion. It is well to be alert to this potential problem and to be clear about which concepts terms refer to in any given context.

The concepts of encoding and retrieval processes, and memory systems, are the most fundamental hypothetical constructs in theory of memory. Brown and Craik (2000, p. 93) explained encoding and retrieval thus:

> The terms *encoding* and *retrieval* have their origins in the information-processing framework of the 1960s, which characterized the human mind/brain as an information-processing device ... In this model, the mind – like the computer – receives informational input that it retains for a variable duration and subsequently outputs in some meaningful form. *Encoding*, therefore, refers to the process of acquiring information or placing it into memory, whereas *retrieval* refers to the process of recovering previously encoded information.

Though the distinction between encoding and retrieval seems relatively straightforward, it is less clearcut than it seems. Encoding entails retrieval. Retrieval entails encoding. The way new events are encoded is heavily dependent on previous experiences, the retrieval of which determines how the new events are perceived and interpreted. Subsequent retrieval of those events in itself creates new events and experiences, which are in turn encoded. Encoding and retrieval are continually interchangeable processes.

The definition of a memory system is more complex. Tulving (1985, pp. 386–387: see also Tulving, 2002) defined memory systems thus:

> Memory systems are organized structures of more elementary operating components. An operating component of a system consists of a neural substrate and its behavioural or cognitive correlates. Some components are shared by all systems, others are shared by only some, and still others are unique to individual systems. Different learning and memory situations involve different concatenations of components from one or more systems ...

Memory systems tend to be defined by a set of criteria, such as differences in the kinds of information they process, in their rules of operation,

some incompatibility in their evolved function, and in the conscious states they may give rise to (Sherry & Schacter, 1987). The biological concept of system offers a useful analogy. For example, organisms have digestive systems, cardiovascular systems, respiratory systems, excretory systems and reproductive systems, and these systems may have common properties and shared components, as well as unique properties and distinct components. And all these systems have certain specific functions, an evolutionary and developmental history, and, of course, physiological and anatomical substrates. Memory systems have much the same general characteristics.

For the last thirty years or so, theory of memory has been divided between theories based exclusively on the processing concepts of encoding and retrieval and theories based on the concept of memory systems. The contrast between these two approaches has generated a great deal of controversy over the years, but more recently there has been some rapprochement between them and an increasing recognition that the two approaches are complementary. Different memory systems all entail encoding and retrieval processes, some of which (or some components of which) they may have in common, some not.

The controversy between these two approaches followed the gradual abandonment of the belief that differences between short-term and long-term memory performance could be explained by a theoretical distinction between short-term and long-term memory stores (Atkinson & Shiffrin, 1968; Waugh & Norman, 1965), a theory that had been so generally accepted that it had become known as the *modal* model of memory. On the one hand, this theory was challenged by the levels-of-processing framework introduced by Craik and Lockhart (1972). On the other hand, this theory was challenged by a further fractionation of memory into additional short- and long-term memory systems, including the working memory model introduced by Baddeley and Hitch (1974) and the distinction between episodic and semantic memory systems introduced by Tulving (1983).

Process theorists have sometimes challenged the concept of memory systems on the grounds that there are no generally agreed 'rules' for proposing the existence of any new system, with the consequent danger of an undesirable proliferation of systems. Systems theorists have responded by suggesting that various criteria, taken together, might help reduce this risk. Process theorists might also be criticized on similar grounds, however, as there are also no generally agreed rules for proposing the existence of any new processes. It can equally well be argued that there has also been an undesirable proliferation of encoding and retrieval processes. But such is the rapid development of the field that no doubt

many of the systems and processes that now seem central may soon be superseded by others, yet to be conceived.

Whatever their ultimate fate, however, current distinctions between different memory systems and memory processes most certainly have heuristic value in investigating memory in special populations such as that of individuals with ASD. These distinctions make the investigation of possible population differences in memory function more tractable. Instead of a global approach in which, perhaps, memory is conceived as a single, undifferentiated entity with memory performance determined largely by stronger or weaker 'trace strength', these distinctions encourage a finer-grained, qualitative approach. Hence, possible population differences in memory function may be found in some memory systems but not others, or in some memory processes but not others. And any differences in memory function that are found can be readily interpreted within an existing body of theoretical knowledge.

The remainder of this introductory chapter is intended to provide a guide to some of those memory systems and processes likely to be of the most immediate relevance to furthering our understanding of memory in ASD. It continues in the next section with a review of major memory systems. This is followed by a section that reviews several key process distinctions. In conclusion, some broader theoretical issues are discussed, including the importance of considering the nature of the memory tasks and of having convergent sources of evidence.

Memory systems

The five memory systems listed in Table 1.1 were identified as such by Schacter and Tulving (1994). Perceptual representation systems are those involved in the perception of objects and events and which represent their structure and form. They give rise to perceptual priming in tasks such as the perceptual identification of objects or of words. Procedural memory refers to those systems involved in skilled behaviour and action and it is usually acquired through extensive practice. Neither

Table 1.1. *Memory systems*

Perceptual representation
Procedural memory
Working memory
Semantic memory
Episodic memory

perceptual nor procedural systems are generally thought to be open to consciousness, in contrast with working memory, semantic memory and episodic memory, where consciousness has a crucial and (arguably) different role in each case. The distinction between these five systems embraces other similar distinctions that include the distinction between short-term and long-term memory (Atkinson & Shiffrin, 1968; Waugh & Norman, 1965) and the distinction between nondeclarative and declarative memory (Squire, 1987). Working memory refers to short-term memory whereas procedural, semantic and episodic memory systems all refer to long-term memory. Semantic and episodic memory systems both refer to declarative memory, whereas procedural and perceptual systems are nondeclarative.

The original working memory model was introduced by Baddeley and Hitch (1974) to replace the unitary view of short-term memory that had characterized the distinction between short-term and long-term memory stores (Atkinson & Shiffrin, 1968; Waugh & Norman, 1965). It has three components concerned with the temporary storage and manipulation of information, a central executive, a phonological loop and a visuo-spatial scratchpad. The central executive – the least well understood component of the model – is viewed as an attentional, supervisory system thought to co-ordinate the operation of the other two components. The phonological loop is involved with the maintenance of phonological information and is crucial for language learning. The visuo-spatial scratchpad is involved with the maintenance of visual and spatial information and is crucial for imagery. The major advantage of this model over the earlier unitary view of short-term memory is that it allows a finer-grained analysis of the functions of short-term memory, including the possibility of selective impairments among those component functions under different task conditions and in different populations.

Working memory depends on its interface with long-term memory systems, both in the retrieval of information from those systems and the encoding of information into them. Baddeley (2000, 2001) has recently introduced an additional component, the episodic buffer. This new subsystem provides temporary storage for the integration of information from other slave systems with information from long-term memory systems. The key point is that the combination of information from different sources itself requires some temporary holding mechanism to bind it together.

Semantic memory is the long-term memory system for general knowledge about the world. It includes information about language; about historical and geographical facts; about music, games, current affairs, and so on. It represents categorical knowledge about concepts. Semantic

memory theories are concerned with how this categorical information is acquired, represented and retrieved. An early model of semantic memory assumed a network of nodes organized in a hierarchy, each node in which representing a discrete concept with links between nodes representing associations among them (Collins & Quillian, 1969). Encoding and retrieval in semantic memory were conceived in terms of spreading activation of nodes and of the connecting links. There are also feature comparison models in which the meaning of a concept is represented by semantic features that may be more or less defining of the concept (Rosch, 1975; Smith, Shoben & Rips, 1974). Later theories include connectionist, or neural network, models (McClelland & Rumelhart, 1985), one fundamental tenet of which is that representation is distributed across the network, in a pattern of activation, rather than being isolated in separate nodes.

Other key concepts in semantic memory are those of schema (Bartlett, 1932) and script (Schank & Abelson, 1977), both of which refer to sets of ideas relating to particular kinds of things such as what classical music sounds like compared with jazz and what clothes are worn in winter, or particular situations such as telling a story or going to the airport to catch a flight. Schemas and scripts represent generalized scenarios reflecting what has been learned about the way the world works and they contain much more information than any simple concept represented by a node in a semantic network. They have an important social role, both in the interpretation of events and in the planning and achievement of goals.

Episodic memory is the long-term memory system for personally experienced events, usually including information about where the events took place and when they occurred. Not all theorists have accepted the need to distinguish episodic from semantic systems, which are clearly closely related. Indeed, it is assumed that episodic memory is built on top of earlier systems, including semantic memory. The most critical feature that separates the two systems is the kind of consciousness experienced when retrieving information from either of them. Retrieval from semantic memory is accompanied by *noetic* awareness, which refers to a sense of knowing, whereas retrieval from episodic memory is accompanied by *autonoetic* awareness, which refers to recollective experiences that entail mentally reliving what was experienced at the time of the original event (Tulving, 1983; 1985). This sense of self in subjective time, or 'mental time travel' as it has been called, has become of increasing importance to the concept of episodic memory (Tulving, 2002) not merely in distinguishing it from semantic memory, but also with respect to its role in thinking about the future. Autonoetic awareness also enables one to

project oneself into the future and it is crucial for the planning of, and for foreseeing the consequences of, personal decisions and actions.

It is important to appreciate that events are registered in semantic memory, as well as in episodic memory, and that according to Tulving's (1995) model of relations between these systems, encoding is serial, storage is parallel, and retrieval is independent. Thus, the occurrence of an event may be encoded in semantic memory without being encoded into episodic memory, but not vice versa. One can know that an event has occurred before without experiencing mental time travel with respect to its previous occurrence. One can know of many previous visits to Paris without re-experiencing anything that happened during any such visit. Semantic memory includes information about one's personal history that is known in a detached and factual way, without the experience of mental time travel. It is also assumed that episodic memory evolved more recently, and develops later in childhood, than semantic memory – perhaps at about the same time as theory of mind (Perner & Ruffman, 1995).

Encoding and retrieval processes

Six process distinctions are listed in Table 1.2. This selection of processes is inevitably more arbitrary than the selection of memory systems, but it does include some of those likely to be useful for investigating memory in ASD. All six distinctions are cast in the form of a dichotomy, though some of them have nonetheless been conceived more as a continuum of processing than as discontinuous categories. For process theorists who have often argued against a systems approach, memory is better approached as a unitary 'faculty' that can be explained in terms of a few broad descriptive and functional principles, such as these, rather than by partitioning it into separate memory systems.

Craik and Lockhart (1972) introduced the levels-of-processing approach as an alternative to theories that distinguished short-term from long-term memory stores (Atkinson & Shiffrin, 1968; Waugh & Norman, 1965). They proposed that memory is simply the by-product of

Table 1.2. *Encoding and retrieval processes*

Deep vs shallow level
Item-specific vs relational
Explicit vs implicit
Conceptual vs perceptual
Effortful vs automatic
Recollection vs familiarity

perception and that deeper levels of processing, in the sense of more meaningful semantic processing, makes for stronger, more durable memories than shallower, more superficial levels of processing. Information can also be maintained at any given level of processing, in short-term memory, but maintenance, *per se*, does not increase the durability of the memory trace. Only deeper, more meaningful processing can do this. This simple set of theoretical ideas has been hugely influential, despite some obvious limitations and shortcomings.

One limitation was that the approach was restricted to encoding. It did not include retrieval. Yet the effects of level of processing at encoding depend on retrieval conditions. Shallow levels of processing may give rise to superior memory performance if the overlap between retrieval and encoding conditions is greater for that level of processing than is the overlap for deeper levels of processing as, for example, when a test requires the retrieval of superficial stimulus features instead of semantic features (Morris, Bransford & Franks, 1977). Such evidence led to the formulation of another important principle, that of 'transfer appropriate processing'. According to this principle, memory performance depends on the extent to which the kind of processing engaged at encoding matches or overlaps with the kind of processing engaged at retrieval. The transfer appropriate processing principle is similar to the encoding specificity principle (Tulving & Thomson, 1973). Encoding specificity was formulated at the level of individual items. It states that no retrieval cue, however strongly related to its target in semantic memory, will aid episodic retrieval unless the information it provides was specifically encoded at the time of study. Transfer appropriate processing is encoding specificity writ large, at the level of the task as a whole, and of the kinds of processing induced by the task.

Levels of processing focuses on the encoding of specific items, and deeper levels of processing in the encoding of specific items makes those items more distinctive. The concept of distinctiveness is a relative concept, in the sense that distinctiveness depends on the context. What is distinctive in one context may not be distinctive in another context. Distinctiveness, like levels of processing, refers to item-specific encoding. It therefore ignores another important concept that had been the focus of much previous research, that of organization. Organization refers to groupings and relations among studied items, and the development of organization during study can greatly increase memory for those items (Bower, 1970; Mandler, 1967). The distinction between item-specific and relational processing usefully embraces both the item-centred focus of levels of processing and the relational focus of organization. In practice, there is often a trade-off between the two. Experimental conditions, or

individual biases, may foster greater relational encoding at the cost of reduced item-specific encoding, or greater item-specific encoding at the cost of reduced relational encoding (Hunt & McDaniel, 1993; Hunt & Seta, 1984).

The use of the terms explicit and implicit memory to refer to different memory systems has now been largely discredited, partly because of the conceptual confusion these terms engendered and partly because of more recent theoretical developments, such as the distinction between semantic and episodic systems, both of which are explicit in the sense that they are both open to consciousness. In contrast with explicit processes, implicit processes are not open to consciousness. Thus, implicit processes refer to the nonconscious forms of memory that are assumed to be reflected in implicit memory tests. But people may often be well aware that they are retrieving studied items in implicit tests, even if they did not intend to retrieve studied items. In view of this, it has been suggested that it is retrieval intention that is critical to comparisons between explicit and implicit tests, rather than awareness that retrieved items were encountered earlier (Richardson-Klavehn *et al.*, 1994; Schacter, Bowers & Booker, 1989).

Perceptual fluency is one implicit process that has been of some theoretical importance in shaping attributional views of memory (Jacoby, 1988; Jacoby, Kelley & Dywan, 1989). Perceptual fluency refers to the perceptual facilitation – some item is perceived more quickly, or more readily if in a degraded form – following a prior act of perception. It has been argued that perceptual fluency gives rise to priming effects in tasks like perceptual identification. Moreover, in recognition memory, the effects of perceptual fluency may be attributed to having encountered the test item in a previously studied list, in the absence of any awareness of the actual occurrence of the item there. Thus, memory is inferred from some other experience. In other circumstances, implicit processes may drive the perception of the stimulus. Thus, a word heard recently may sound louder next time than one not heard recently, or the previously studied name of a nonfamous person may seem famous, when making fame judgements about those names in the context of names of other moderately famous people. An attributional view of memory is concerned with how memory may be inferred from other kinds of experiences and with how memory may influence other kinds of experiences (Jacoby, 1988; Jacoby, Kelley & Dywan, 1989).

The distinction between conceptual and perceptual processes was largely developed in order to provide an alternative account of dissociations between memory performance in explicit and implicit tests to the account provided by the theory that the two kinds of tests involve