A **Routledge** FREEBOOK

Memory and Cognition

Topics in Psychology

PSYPAG Psychology Postgraduate Affairs Group

Routledge Taylor & Francis Group

TABLE OF CONTENTS

About PsyPAG

PsyPAG is a national organisation for all psychology postgraduates based at UK Institutions. Funded by the Research Board of the British Psychological Society, PsyPAG is run on a voluntary basis by postgraduates for postgraduates.

PsyPAG aims to:

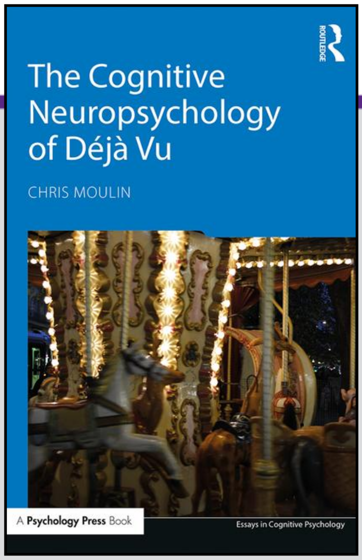
1. Provide support for postgraduate students in the UK.
2. Act as a vehicle for communication between postgraduates, and to represent postgraduates within the British Psychological Society (BPS).
3. Bring together postgraduates from across the country.

PsyPAG has no official membership scheme; anyone involved in postgraduate study in psychology at a UK Institution is automatically a member!

Find out more at http://www.psypag.co.uk.

CHAPTER 1

THE HUMAN RECOGNITION MEMORY SYSTEM



This chapter is excerpted from
The Cognitive Neuropsychology of Déjà Vu
by Chris Moulin.

Learn more

THE HUMAN RECOGNITION MEMORY SYSTEM

Excerpted from *The Cognitive Neuropsychology of Déjà Vu*

> My memory is extensive, yet hazy: it suffices to make me cautious by vaguely telling me that I have observed or read something as opposed to the conclusion which I am drawing, or on the other hand in favour of it; and after a time I can generally recollect where to search for my authority.
>
> *Charles Darwin (1887)*

The central thesis of this book is that the phenomenon of déjà vu can be considered within existing theories of human recognition memory. Recognition is the memory system that is responsible for detecting prior occurrences of stimuli in the environment. In short, déjà vu presumably arises because of a temporary glitch or misinterpretation of the recognition memory system, a system that is usually responsible for responding appropriately to familiar and novel environments, ideas and people. In this chapter, an overview of the recognition memory system is presented, focusing in particular on two key concepts: recollection and familiarity. The purpose is to give a neuropsychological account of how déjà vu may be considered in terms of the healthy function of the memory system. The chapter has two parts. The first gives an overview of recollection and familiarity. The second introduces the concept of epistemic feelings that are thought to govern memory function.

Dual processes theories of recognition memory

Recognition memory rests on a decision-making process. When we detect something as having been encountered before we make a comparison between what is represented in the cognitive system and what is currently perceived.When encountering a stimulus

THE HUMAN RECOGNITION MEMORY SYSTEM

Excerpted from *The Cognitive Neuropsychology of Déjà Vu*

(such as a newly learned word, for example) for a second time, a number of processes and sources of information are brought to bear on your processing of the word: how fluently you can process it, the distinctiveness of its perceptual trace, the feelings generated when encountering it a second time, the effort involved in retrieving its meaning, and whether you can recall the specifics of your first encounter with the word: who used it and in what context. These complex sources of information can be used to retrieve the meaning of the word, or to gauge the certainty with which you have encountered the word before, and so on. Ultimately, they can all be used to make a decision about whether we have encountered the word before or not.

Recent theories of human decision making suggest that complex tasks requiring problem solving and judgement rely on two different types of thinking (Evans, 2008; Kahneman, 2011). Such dual-process accounts consider that people make decisions based on two separable streams of information: a fast, intuitive feeling and a slower, more deliberative evaluation. These separable processes in cognition are arguably also at play in memory decision making (e.g. Arango-Muñoz, 2010; Hintzman & Curran, 1994; Koriat & Levy-Sadot, 2001) and map neatly onto the concepts of familiarity (which is relatively fast and automatic) and recollection (which is slower and more deliberative) (Mandler, 2008). Human recognition memory decision making can be thought of as the combination of information drawn from two different processes: recollection and familiarity.

The idea of dual processes in memory and cognition is not new, and was probably inspired by the anatomy of the brain. Many early scholars posited that the two hemispheres of the brain represented a 'double organ' (e.g. Holland, 1840).Wigan's influential text (1844), *The Duality of the Brain*, was an extreme position. He argued

THE HUMAN RECOGNITION MEMORY SYSTEM

that there were literally two separate brains that could work in synchrony or not. This early view of the 'dual' brain influenced Ribot's early conceptions of human memory and its disorders (Taylor & Shuttleworth, 1998). One prominent view of déjà vu is that it arises from a mismatch between two separate streams of consciousness (often blamed upon two hemispheres working out of synchronisation), something that Wigan discussed as early as 1844.

The contemporary view of recognition memory is that it is based on neurally distinct mechanisms of recollection and familiarity. Yonelinas (2002) describes familiarity as a direct evaluation of the memory trace, something that can be subjectively reported by the participant. In contrast, recollection refers to the retrieval of specific contextual information from the time of study, and is often characterised as 'mental time travel' or as having the first-person experience of 'remembering'. Recollection has been characterised as the ability to recall or report 'something more' at the time of making a recognition decision (Moulin, Souchay, & Morris, 2013). This can be captured by asking participants to justify their responses – "SUSHI – I know I've seen this word before because I remember thinking about my cat while it was presented earlier." Equally, it can be measured by a source-memory task (what colour ink a word was presented in, whether it was a male or female voice that spoke it, etc.).

The dual-process view of recognition memory is contentious, and more parsimonious single-process theories exist (see Diana, Reder,Arndt, & Park, 2006; Donaldson, 1996; Dunn, 2004; Squire, Wixted, & Clark, 2007; and Wixted & Stretch, 2004). Those who argue that remembering and familiarity lie on a continuum and reflect just one memory system suggest that the two differ along one dimension, the strength of the trace in memory. In experimentation, this is usually captured empirically as subjective confidence. Participants report higher levels of confidence for events that are 'remembered' than for those which merely feel familiar (Dunn, 2004). Recently, Berry, Shanks, Speekenbrink and Henson (2012) have argued that even implicit and explicit memory phenomena lie along the same continuum. In a series of experiments, they show that the rate at which someone can identify a very briefly presented word (typically seen as an implicit memory test) is related to the subjective report of whether the word has previously been seen. Critical for their argument, a mathematical model with one 'process' reproduces their data. In short, the same underlying trace strength is supposed to support not only familiarity and recollection decisions, but even implicit and explicit memory phenomena – priming and recognition memory.

The temporal lobe memory system

Neuroscientific and anatomical data can help resolve the single/dual-process argument, and the neural basis of recollection and familiarity is currently under debate. Aggleton and Brown (1999) put forward a widely cited neuroscientific model, focusing on the hippocampus as critical for recollection, and the adjacent parahip-pocampal gyrus as responsible for familiarity. Their proposal is briefly summarised in Figure 3.1. Aggleton and Brown further suggested that due to the network connecting the hippocampus to the fornix, mamillary bodies and anterior thalamic nuclei, these structures are also engaged during the encoding and retrieval stages of recollection. Moreover, they suggested that familiarity is supported specifically by the most anterior portion of the parahippocampal region (the perirhinal cortex, PRc).

THE HUMAN RECOGNITION MEMORY SYSTEM

Excerpted from *The Cognitive Neuropsychology of Déjà Vu*

This model predicts that hippocampal damage should affect recollection but not familiarity, and parahippocampal damage should lead to impairments in familiarity, not recollection. In support, patients with damage restricted to the hippocampus have displayed isolated impairments in recollection measured through a number of paradigms (e.g. Bowles et al., 2010). Later models have elaborated on the specific roles of the PRc, entorhinal (ERc) and parahippocampal cortices (PHc) due to the emergence of findings that extrahippocampal structures may be able to support some forms of associative memory. Such a departure, as Montaldi and Mayes (2010) describe, begins to view recollection and familiarity as 'kinds' of memory, because "each is a complex function, likely to depend on several different processes that are probably mediated by different structures that are functionally connected in a system" (p. 1294).

A meta analysis of event-related functional magnetic resonance imaging (fMRI) studies on healthy participants by Diana, Yonelinas and Ranganath (2007) found that recollection was associated with relatively more activation in the hippocampus than the perirhinal cortex, whereas familiarity was associated with relatively more activation in the perirhinal cortex than the hippocampus. However, they also argue that "there is no simple mapping between MTL regions and recollection and familiarity, but rather that the involvement of MTL regions in these processes depends on the specific demands of the task and the type of information involved" (p. 379).

Just as cognitive single-trace accounts contest the assumptions of dual-process theories, there is also opposition to the above neuroanatomical models. Squire and colleagues (e.g. Squire et al., 2007; Wixted & Squire, 2011) argue that all structures within the medial temporal lobe (MTL) mediate recollection and familiarity equally.

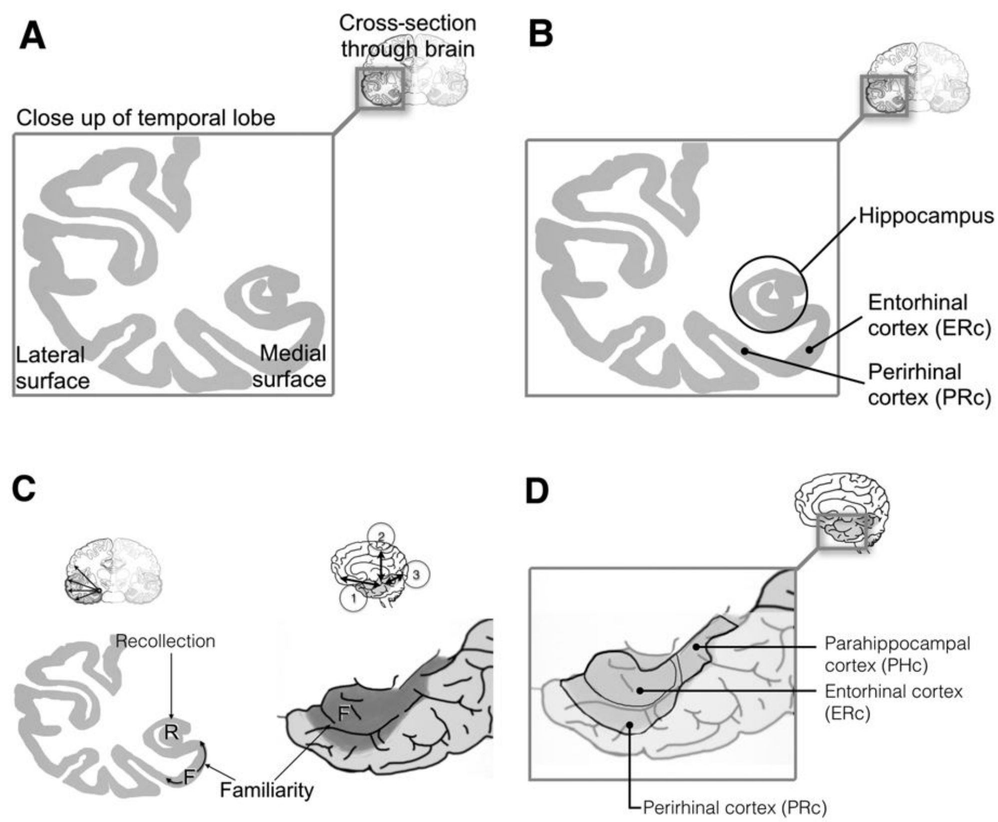


FIGURE 3.1 Anatomical schematic of recollection and familiarity. (A) Overview. (B) Coronal view. (C) Crude depiction of Recollection and Familiarity. (D) Sagittal view.

Their MTL Unitary Trace Strength (MUST) account suggests that functional heterogeneity exists within the MTL, but not for recollection and familiarity. Wixted argues (e.g. Wixted, 2007; Wixted & Squire, 2011) that recollection is a continuous process just like familiarity. In a source-memory experiment using fMRI, Wais, Squire and Wixted (2010) measured hippocampal activity at retrieval after equating memory strength of recognition decisions on item-correct plus source-correct or item-correct plus source-incorrect trials. Their analysis focused only on recognition trials where participants assigned high confidence ratings, regardless of whether the correct source was retrieved. They found that hippocampal activity was similarly elevated for both correct and incorrect source judgements, suggesting it is involved in both recollection and familiarity.

It is crude to think of the hippocampus and parahippocampus as the only areas of the brain responsible for recognition memory, even though disease and damage of these areas impairs memory function. fMRI investigations of episodic memory reveal that brain regions associated with attention and decision making are more reliably activated during memory tasks than would be expected compared to findings from patients with brain damage. In particular, the parietal and prefrontal cortices are reliably activated during episodic memory tasks. The parietal lobe is often thought of as integrating sensory information. In healthy groups, parietal lobe activation has been attributed to the support of episodic recollection at retrieval (as with the robust EEG finding of a P600 neural signature for recollection) and is involved in hippocampocortical memory and fronto-parietal 'resting state' networks (for a review of parietal lobe involvement in memory, see Wagner, Shannon, Kahn, & Buckner, 2005). That is, a large network including the parietal and frontal regions possibly determines memory activity and coordinates the whole memory network.

Frontal structures have long been thought of as responsible for controlling and monitoring the cognitive system (e.g. Stuss & Levine, 2002). They are seen as coordinating, reflecting upon and controlling the temporal lobe memory system (Moscovitch, 1994). The idea of the 'controlling' function of the frontal lobes is manifest in the neuropsychological deficits following frontal lobe damage, most notably confabulation (e.g. Burgess & Shallice,

THE HUMAN RECOGNITION MEMORY SYSTEM

Excerpted from *The Cognitive Neuropsychology of Déjà Vu*

1996). Fletcher and Henson (2001) reviewed the role of the frontal lobes in human memory as shown in neuroimaging studies. They subdivide the frontal lobes into three critical zones: the lateral surfaces of the ventrolateral, dorsolateral and anterior frontal cortex. They propose that these respective regions are responsible for updating/maintenance of information; selection/ manipulation/ monitoring of information; and the selection of processes and subgoals. Fletcher and Henson point out that whereas frontal lobe damage is very disruptive to working memory, it has a more restricted effect upon long-term memory. Patients with frontal lobe damage fail to suppress interference (e.g. della Rocchetta & Milner, 1993) and cannot reproduce temporal order (Shimamura, Janowsky, & Squire, 1990). Stuss and Levine (2002) also emphasise that frontal lobe damage leads to difficulties in organising material during retrieval and effectively generating cues to reproduce information. Frontal lobe damage contributes to, but does not in itself cause, a global amnesia.

Neuroimaging studies point to the activation of a 'retrieval mode' in the frontal lobe. In a recognition task where participants had to either respond 'yes' to all words they had seen previously, or to only reply 'yes' to words seen in a particular context, Henson, Rugg, Shallice, Josephs and Dolan (1999) found activation bilaterally in the dorsolateral frontal cortex, which is interpreted as reflecting source monitoring of the retrieved words (see also Mitchell & Johnson, 2009, for similar findings). Fletcher and Henson conclude that retrieval from episodic memory includes two main stages of frontal lobe involvement: the generation of search parameters, and the verification of the material generated in a memory search – and they propose these processes map onto ventrolateral and dorsolateral cortices, respectively.

Perhaps most critically for the study of déjà vu, patients with damage to the frontal lobes have been known to develop problems with false recognition. Schacter, Curran, Galluccio, Milberg, and Bates (1996) describe patient BG, who following an infarction of the right frontal lobe had abnormally high levels of false recognition accompanying the subjective experience of remembering (see also Curran, Schacter, Norman, & Galluccio, 1997). BG not only incorrectly recognised items (words, pictures and sounds) he had not previously studied, but he also reported 'remembering' these items, rather than just finding them familiar. BG did not show this pattern when required to learn a categorised list. Schacter et al. (1996) concluded that BG was only aware, when tested on uncategorised lists, that he had previously studied a list consisting of some words, pictures or sounds. When presented with a recognition test containing old and new items, he over-extended his 'recognition' to new items, perhaps on the inference that as these were similar to the previously studied items they were likely to be old. However, if more structure was present in the encoding environment, for instance from semantic categories, then he was able to use this structure to discriminate old from new items. Such cases of frontal impairment support the idea that the frontal lobes are involved in the coordination and monitoring of memory processes.

In sum, whereas there is consensus that recollection and familiarity are distinguishable psychological concepts, their status as neurological and/or cognitively separate entities is in question. Given that the two responsible brain regions – the hippocampus and the parahippocampus are both tiny and boast many complex inter-connections between themselves and other brain areas, it is likely that a more nuanced view of these zones is required. Moreover, one needs to consider that these regions are imbedded into a large network of regions all implicated in memory function.

THE HUMAN RECOGNITION MEMORY SYSTEM

Excerpted from *The Cognitive Neuropsychology of Déjà Vu*

What is familiarity?

Familiarity is the central subjective experience in déjà vu. Assuming that this form of familiarity is that which is found in recognition memory, it is important to describe in detail how this experience is generated in the memory system. One problem with discussing familiarity is that – just as with déjà vu – there exist alter- native definitions of this technical term. One of the most influential theories of familiarity comes from Bruce Whittlesea (e.g. Whittlesea, 1997). He pointed out that there are three common uses for the term familiarity that contribute to some confusion in the area:

1. "... a person has actually encountered a stimulus (or even one like it) previously.This sense pertains to the historical fact that a person has previous experience with an object, whether or not that experience influences current behavior and whether or not the person can report that experience."

2. "... the person has knowledge about a stimulus that permits them to perform appropriately toward an object, without necessarily having an accompanying feeling of having experienced that stimulus previously. For example, in watching Hamlet for the 15th time, I know what to expect is coming next, but I have no pressing feeling of having seen it before."

3. "... the subjective feeling of having encountered a stimulus on some previous occasion, whether one actually has or not."

Whittlesea & Williams (1998, pp. 141–142)

Whereas cognitive psychologists may use familiarity in all these senses, for our purposes we are clearly interested in *subjective* familiarity, which is emphasised in the third definition. This definition captures the idea of subjectivity critical to the feeling of déjà vu – it is also, importantly, suggested in this definition that this form of familiarity can actually be false, which is both critical for our definition of déjà vu as being false, and as will be developed below, is also important for experimentation in human memory and inducing familiarity.

The butcher on the bus phenomenon

Subjective familiarity is perhaps most strongly felt when we are unable to know why we feel it so strongly, as Mandler discussed in 1980:

> Consider seeing a man on a bus whom you are sure that you have seen before; you "know" him in that sense. Such a recognition is usually followed by a search process asking, in effect,Where could I know him from? Who is he? The search process generates likely contexts (Do I know him from work; is he a movie star, a TV commentator, the milkman?). Eventually the search may end with the insight,That's the butcher from the supermarket!
>
> *Mandler (1980, pp. 252–253)*

This is the 'butcher on the bus' phenomenon. It describes the frustrating experience that occurs when we are unable to pinpoint the source of the familiarity. It shows that familiarity can be felt especially strongly in the absence of certain knowledge of how the person is familiar. Familiarity or 'just knowing' is the simple judgement of prior occurrence, and at its simplest, familiarity is easy to define, because it is a judgement of prior occurrence that is devoid of the retrieval of contextual specifics. Note that this sense of familiarity, although intensely subjective, is not the same as déjà vu. Somehow in the butcher on the bus phenomenon, we are sure that we are searching for a reason why the feeling is true – a search for why the person feels familiar. In comparison, in the déjà vu experience we know – immediately or soon after – that the strong sense of familiarity is false.

This does little to answer what familiarity *is.* Yonelinas (e.g. Yonelinas, 2002) sees familiarity as a 'signal-detection process'. A signal-detection process account suggests that at high levels of familiarity the trace strength of the memory is high, and we are able to read off the strength of the memory signal in order to report,'I have seen this face before'. For weaker activations, familiarity is not sufficient to make a conscious report of prior occurrence, but it might be sufficient to influence our preferences, or 'gut' feelings. Thus, familiarity is a graded process with the strength of familiarity for a stimulus influencing behaviour. An item can be familiar, but below the level of conscious report.

Whittlesea's view is that the memory system is continuously trying to make sense of its inputs, so that it can interpret any signals arising from low-level processing of the environment. In this way, familiarity is a subjective feeling arising from the fluent processing of a stimulus. It is not an inherent feature of anything you have seen before. Familiarity is an attribution we can make to explain why we come to know something quickly, or why we can fluently process a word or a face. His theory is summarised in the title of one of his papers: 'Why do strangers feel familiar, but friends don't?' (Whittlesea & Williams, 1998). We do not feel overwhelming familiarity when we come home and see our husbands: we were expecting to see them there in the house at that time. His explanation of the butcher on the bus is that it is the fact that we cannot retrieve who the person is that gives rise to the feeling of familiarity. This theory is supported by various illusions of familiarity: where we use the fluency of processing to wrongly judge that we have seen something before. Goldinger and Hansen (2005) report an experiment for which they built a chair that administered subliminal buzzes through the seat. They were interested in the way in which these subliminal buzzes influence recognition memory decision making. Participants sat on the chair and then conducted an episodic recognition memory test, where they reported 'old' or 'new' for a set of words, and also reported the confidence in their answer. The results showed that the subliminal buzz influenced the recognition decision (particularly for the most difficult items on the test). The effect was to increase the rate of responding 'old' to the items on the test, both the targets and the distracters in the recognition test. It also increased the confidence made to the false alarms, but reduced it for the hits. On his website, Goldinger describes his results thus:

> When you truly have no memory, the buzz gives you a tingle of confidence, but when you have a memory, the buzz gives you a tingle of doubt. These findings were in line with the predictions of a model called SCAPE, as the same signal created a different memorial interpretation, based on context (Whittlesea & Williams, 2001).
>
> *Goldinger (2017)*

These results suggest that low-level feelings are used when making explicit decisions about memory – it suggests an interchange between feelings and thoughts, and non-conscious feelings and explicit decisions. Whittlesea's SCAPE (Selective

THE HUMAN RECOGNITION MEMORY SYSTEM

Excerpted from *The Cognitive Neuropsychology of Déjà Vu*

Construction and Preservation of Experience) model (Whittlesea & Williams, 2001) describes the interaction of attributions of memory and processing fluency. The current context generates a top-down expectation of processing fluency.When an expectation of fluency is violated, it triggers a search for why that has arisen. This is the search triggered in the case of the butcher on the bus. In the case of the buzz chair, the buzz contributes to the decision (note that if the buzz is not subliminal, the same effect does not occur): the attribution of the buzz is that it must 'mean' something for the ongoing memory task.

These experimental results converge on the idea that familiarity is a feeling that is generated from the processing of a stimulus – if a stimulus is processed very fluently, our attribution is that we have encountered it before. For a philosophical account of this possible 'attributionalist' nature of episodic memory, see Perrin and Rousset (2014). The fact that we can alter subjective familiarity and recognition memory decision making by altering the fluency with which items are processed suggests that we use familiarity signals in making attributions about whether we have encountered information before. As we shall see in Chapters 5 and 10, this means that generating false feelings of familiarity should be sufficient for generating déjà vu in the laboratory.

Neuroscientifically, familiarity can be seen as a low-level, quick-acting form of memory which works on perceptual inputs. Similarly, this system tries to match, as quickly and effortlessly as possible, the contents of mental representations stored in memory with the current contents of perception. Viewed like this, familiarity decision making is the last stage in perception – once we have composed and identified a scene or environment, we can 'read off' whether we have encountered it before. This is consistent with a neuroanatomical view of the temporal lobe memory system being the last point in the ventral visual stream (e.g. Suzuki & Amaral, 1994).

The remember/know paradigm

The dual-process debate is most pronounced when comparing the subjective experiences accompanying recognition memory in the remember/know paradigm, which has been used throughout the literature to examine episodic recognition memory (e.g. Tulving, 1985). The remember/know paradigm asks participants to distinguish between sensations of 'remembering' and 'knowing' (and/or finding familiar, e.g. Conway, Gardiner, Perfect, Anderson, & Cohen, 1997). Common definitions of these concepts are presented in Table 3.1. The paradigm is a straightforward means of estimating the relative contributions of the two separate processes in recognition memory. There are other means of assessing recollection and familiarity, as will be shown in later chapters, but these other methods tend to converge on the same construct (Eichenbaum, Yonelinas, & Ranganath, 2007).

The critical issue with the remember/know paradigm is the evidence for separable streams of familiarity and recollection in recognition memory. Importantly, Perfect, Mayes, Downes and Van Eijk (1996) tested both subjective report of remembering and the recall of contextual (or source) information: familiar responses rarely yielded reports of contextual information (e.g. where on the screen the item was presented) that were above chance. When participants *remembered* items at test, they were more reliably able to report source details from the study phase, suggesting that the subjective feeling of remembering is indeed related to the recall of source information.

THE HUMAN RECOGNITION MEMORY SYSTEM

Excerpted from *The Cognitive Neuropsychology of Déjà Vu*

TABLE 3.1 Definitions of remembering and knowing. Selected quotations detailing subjective experiences were described to participants in various studies (adapted from Williams & Moulin, 2015, p. 983)

Authors	*Response options in experiment*	*Representative quote and/or definitions provided to participants*
Gardiner and Java (1990)	Remember Know	"Often, when *remembering* a previous event or occurrence, we consciously recollect and become aware of aspects of the previous experience. At other times, we simply *know* that something has occurred before, but without being able consciously to recollect anything about its occurrence or what we experienced at the time" (p. 25, emphasis in original)
Bastin and Van der Linden (2003)	Remember Know Guess	"… classify a 'yes' response … as 'Know' if you do not remember any information associated with the face.You are *sure* that you have seen it because you have a *strong feeling* of *familiarity*, but you do not remember any information encoded with the face" (p. 24, emphasis added)
Gardiner, Java, and Richardson-Klavehn (1996)	Remember Know Guess	"The subjects were told that a know response meant that they *knew for a fact* that the word occurred in the study list, because the word was *familiar* in the experimental context, but they did not recollect its occurrence" (p. 116, emphasis added)
Kelley and Jacoby (1998)	Remember Know	"A Know response is defined as the inability to recollect any details of the study presentation in combination with a *feeling of familiarity* or *certainty* that the word was studied" (p. 134, emphasis added)
Dewhurst and Anderson (1999)	Remember Know Guess	"A know response is one in which you recognize the item because it *feels familiar* in this context, but you cannot recall its actual occurrence in the earlier phase of the experiment.You recognize the item *purely on the basis of a feeling of familiarity*" (p. 667, emphasis added)

Moreover, remembering and familiarity are affected differently by experimental factors (reviewed by Yonelinas, 2002). In short, there are a number of neuropsychological dissociations that are important (such as in healthy aging, autism, Alzheimer's disease). There are also a number of manipulations at study and at test that can alter either familiarity or recollection, suggesting that they are separable processes. For instance, during encoding, a deep level of processing compared to a shallow level of processing leads to a greater change in recollection than familiarity.

Epistemic feelings and metacognition

Up to this point, a family of illusions and phenomena have been presented that give an idea of how familiarity works in the recognition memory system and how we can use feelings of familiarity and fluency to make inferences about prior occurrence. At some level these acts of attribution and making meaning of processing are metacognitive (i.e. thinking about thinking) because they involve reflections about our own performance. The remember/know paradigm can be loosely described as metacognitive, because it considers the first-person report of experiences and feelings that are involved in a recognition memory decision. A parallel literature considers such guiding metacognitive phenomena as 'epistemic feelings', which are central to cognitive processes. As such, the feeling of familiarity can be described as an epistemic feeling in that it is a sensation that guides our cognitive behaviours (Moulin & Souchay, 2014). Interestingly, as with human recognition memory, the discussion of epistemic feelings is often in the context of discussing dual-process theories, with epistemic feelings such as familiarity acting as fast, intuitive cues which guide cognitive processing.

Koriat (2007) described the low-level subjective states in memory processing as *experience-based metacognition*. Based on Koriat's

THE HUMAN RECOGNITION MEMORY SYSTEM

view, Arango-Muñoz (2011) suggests that metacognition can be split into two levels, one that is metarepresentational, and can be described as 'thinking about thinking'. This includes making predictions of future performance based on the parameters of the memory task, the characteristics of the to-be-remembered stimuli, and our general dispositional characteristics, such as *I can never remember street names*. This basically concerns humans turning their ability to predict other people's intentions and behaviours on themselves (e.g. Flavell, 2004). Arango-Muñoz discusses a second level of *epistemic feelings* which is a quick-acting intuitive process, based on how things *feel* rather than an assessment based on stored representations and problem-solving heuristics. These include feelings of certainty, pastness, insight, fluency and mental effort (de Sousa, 2009). Again, these two levels of metacognition map nicely onto relatively recent developments in dual-process accounts of reasoning and judgement; distinguishing between cognitive processes that are fast and automatic, and those that are slower and deliberative. Epistemic feelings are proposed to give a sense of 'truth' to a belief, or ascribe some meaning to our cognitive processing. As with emotions more generally, it is crude to think of these feelings as 'right' or'wrong'– there are times when we cannot (*and should not*) objectively say whether someone is 'correct' to feel angry, for instance. However, we can consider, as with emotion, what the experiment considers the feeling is 'about': people can report how the feeling is being interpreted – it is this face, or that word, which feels familiar, for example. For a philosophical debate about the nature of these epistemic feelings, see Proust (2007). Arguably, they are fast-acting, reflective and guide-processing in the same way that emotional feelings are, but much like an emotion, they do not deliver any content – they are just a signal.

THE HUMAN RECOGNITION MEMORY SYSTEM

Excerpted from *The Cognitive Neuropsychology of Déjà Vu*

There are a number of different feelings and states that can be experienced when trying to retrieve information, ranging from fast, automatic 'ecphoric' retrieval of facts to a complete failure to recognise a person or place, such as with the butcher on the bus. In many instances, memory retrieval will be strategic, and information and feelings generated during failed recall will be useful for the experiment and guide-processing. We do not need to search memory in order to make a recognition decision on any one item – in this way there is a difference between failing to retrieve something and knowing that we don't know it. For instance, you can very quickly answer questions like "Did you ever eat in *DZ'envies* in Dijon, France?" without setting up an exhaustive search of French cities and their restaurants (see Kolers & Palef, 1976).There is very little research on how we 'know not', but presumably the capacity to quickly know that something is novel or unknown is one way we can generate the conflict in déjà vu: *this feels familiar but even without questioning my memory I know I've never been here before*.

Thus, an understanding of these feelings and their interpretation is central to the study of déjà vu, and in fact, déjà vu is often used by philosophers as an example of an epistemic feeling: when there is a sensation of memory without the content of memory. Philosophical material relating to déjà vu is beyond the scope of this book, but for some philosophical debates about déjà vu and recollective confabulation (Chapter 8) see de Sousa (2009), Brun, Doğuoğlu and Kuenzle (2009), Arango-Muñoz (2014), Bortolotti (2010) and Gerrans (2014). Moreover, understanding the metacognitive contributions to memory is critical. Roediger (1996) asserts that déjà vu is an "illusion of metacognition" (p. 95). Brown (2004) describes déjà vu as a "'pure' metamemory experience unconnected with the empirical world" (p. 5).

THE HUMAN RECOGNITION MEMORY SYSTEM

Excerpted from *The Cognitive Neuropsychology of Déjà Vu*

The literature on epistemic feelings in memory is rather underdeveloped, but the concept neatly maps onto a set of metacognitive phenomena and paradigms that are rather better researched and understood (such as the feeling of knowing, or FOK; e.g. Wojcik, Moulin, & Souchay, 2013). Epistemic feelings are possibly easiest to describe in instances where there is a mismatch or dissociation between the processing and the contents of cognition, such as in déjà vu, but also in the tip-of-the-tongue (TOT) state. A TOT occurs when there is a feeling that a piece of information is known, but it is not available for conscious report. In TOT and déjà vu there is evidence for the existence of epistemic feelings in that a feeling about cognitive processing has become divorced from the material being processed. Normally, when epistemic feelings are in concert with the goals of processing, we are not so aware of them, just as with Whittlesea's observation about friends not feeling familiar.

Summary: familiarity and recollection in episodic memory

Human recognition memory processes rely on processing in the medial temporal lobe and recognition memory decision making is subserved by two neurologically distinct but adjoining regions which are responsible for recollection and familiarity. In addition, frontal areas of the brain act in a network with the temporal lobes and most likely play a role in interpreting the activations in the temporal lobe. Crudely speaking, it can already be seen that an anomaly in activation in parahippocampal circuits may generate a feeling of false familiarity that is possibly divorced from the usual activation, which sees familiarity triggered as the last part of the ventral visual stream. The exact mechanics of déjà vu still need elucidation, but the temporal lobe is clearly the brain region most likely to be responsible for this experience, both in terms of early attempts at understanding the dreamy state and epilepsy and direct stimulation of the temporal lobe (see previous chapter), but this is also in agreement with contemporary theory about human memory.

In short, déjà vu can be seen as a false feeling of familiarity, which is opposed by top-down contextual information which points to this epistemic feeling being false. Clearly, a central proposal for this thesis is that familiarity can be viewed as a low- level memory 'experience' based on a (intuitive and fast) metacognitive evaluation of the memory system. Souchay and Moulin (2009) and Moulin et al. (2013) have summarised the above theories into a schematic representation about the relationship between familiarity and recollection in recognition memory. This model has been used to explain the search for information given the recognition of a stimulus in the environment, and pertains most clearly to the feeling of knowing (FOK; Hart, 1965). The FOK is a situation where one has the feeling or belief that they will later be able to recognise a currently inaccessible piece of information. Figure 3.2 shows a characterisation of the relationship between familiarity and recollection. This model proposes that the relationship between familiarity and recollection is metacognitive: familiarity is a signal which means something in the human memory system – that is we can act on our feeling of familiarity in order to guide memory search.

In this model, familiarity is proposed to be a trigger for the search for contextual information and occurs before the recollection of specifics (cf. Hintzman & Curran, 1994; Koriat & Levy-Sadot, 2001). An initial feeling of familiarity with a cue triggers the attempted retrieval of episodic detail about the cue and any associated information. In the absence of any contextual information, we can make accurate predictions of our future recognition (as is measured in the episodic FOK paradigm, e.g. Souchay, Moulin, Clarys, Taconnat, & Isingrini, 2007) based on the strength of familiarity for the cue, and also based on what other 'partial information' about the cue or searched-for-target comes to mind. This model is metacognitive because, judging our capacity to correctly gauge correct recognition of a currently unrecallable stimulus, it appears that we can use the familiarity of a cue in memory to make evaluations of the state of a target in memory.

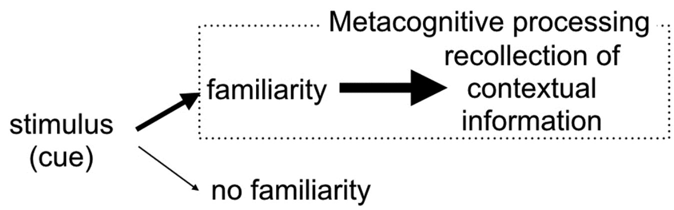


FIGURE 3.2 A sequential model of familiarity and recollection based on an initial feeling of familiarity and a subsequent search of memory. The relationship between familiarity and recollection in this model is metacognitive.

That is, familiarity can be acted upon metacognitively to investigate why – as in the case of the butcher on the bus – we find a stimulus familiar. This situation is analogous to searching for a target word when given a cue word. The model is a little simple, but it predicts, for instance, that we cannot recollect information for something that does not feel familiar to us. It does, however, suggest that, as in the déjà vu experience, people can find something familiar, but not be able to retrieve any information as to why it is familiar. Such a model views déjà vu as metacognitive:

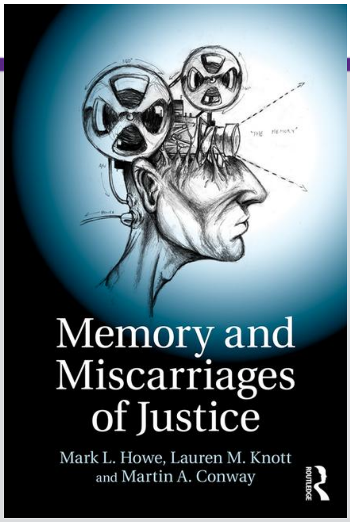
THE HUMAN RECOGNITION MEMORY SYSTEM

familiarity may be erroneously triggered, but then the search for contextual information does not return supporting recollective information. Instead, there is the certainty that the familiarity is false, which is presumably supported by an evaluation that the current situation/location/perceptual input is novel. Note that this is not the same as the butcher on the bus, where familiarity is high but the recollected specifics are (momentarily) absent. It is rather that top-down information is generated which actually opposes the feeling of familiarity.

CHAPTER

2

MYTHS AND NAÏVE BELIEFS ABOUT MEMORY



This chapter is excerpted from

Memory and Miscarriages of Justice

by Mark L. Howe, Lauren M. Knott and Martin A. Conway.

Learn more

MYTHS AND NAÏVE BELIEFS ABOUT MEMORY

Excerpted from *Memories and Miscarriages of Justice*

As we have seen in Chapter 2, eyewitness evidence in criminal cases is one of the most important and most frequently encountered pieces of evidence we see in the courtroom. In a study conducted in 1987 it was estimated that in 77, 000 criminal trials each year in the United States, the primary or sole evidence was testimony provided by eyewitnesses (Wells, Small, Penrod, Malpass, Fulero, & Brimacombe, 1998). Yet we know from the research in this area, and the hard work of the Innocence Project, that eyewitness error is the leading cause of wrongful convictions.

Police officers, lawyers, judges, and jurors regularly face the complex and difficult task of knowing when to accept an eyewitness' testimony as reliable or unreliable depending on the exposure to memory-distorting factors. Similarly, clinicians must consider when a client reveals a memory or personal story within a therapy session, whether this reflects a genuine experience, whether their memories have been distorted by factors associated with time, or whether this memory is one that reflects a fictitious event (Qin, Goodman, Bottoms, & Shaver, 1998). Sorting memories based on fact and memories based on fiction is no easy task, but having knowledge about the fallibility of autobiographical memory and the factors that may lead to the distortion of memory would appear crucial to success (Magnussen & Melinder, 2012). In an ideal world, all those that encounter memory evidence, be it in a legal or clinical setting, should possess sufficient knowledge in order to determine the authenticity of autobiographical recollections. Of course, as it stands now, it appears that many of us (i.e., laypeople, clinicians, police officers, lawyers, jurors, and judges) hold naïve beliefs about how memory operates as well as how we come to decide the reliability of memory when it serves as evidence (Magnussen & Melinder, 2012).

The evidence to support the notion that naïve beliefs about memory are more common than one might anticipate comes from several recent surveys that have probed both the general public (Conway, Justice, & Morrison, 2014; Magnussen et al., 2006) and professional groups including those involved in the judicial system as well as those in clinical settings (e.g., Magnussen & Melinder, 2012; Magnussen et al., 2008; Wise & Safer, 2004; Wise, Safer, & Maro, 2011). The surveys have asked questions about beliefs concerning memory in general as well as about issues relating to the reliability of eyewitness testimony specifically. For example, between 45% (Magnussen et al., 2006) and 81% (Patihis et al., 2014) of jury-eligible lay people believe that frightening and traumatic memories can be blocked or repressed. However, scientific evidence runs strongly against that belief (see Chapter 2; Loftus & Ketcham, 1994). Moreover, 70% of people believe that these repressed memories can be subsequently retrieved accurately in therapy (Patihis et al., 2014), whereas scientific evidence stresses the distortive nature of therapeutic memory recovery techniques (see Chapter 2). Finally, approximately 55% of people believe that memory accuracy can be enhanced by hypnosis (Patihis et al., 2014; Simons & Chabris, 2011), whereas the scientific literature is replete with demonstrations of the highly suggestive qualities of hypnosis (see Chapter 2; Ran, Shapiro, Fan, & Posner, 2002).

Similar findings have been reported by clinical investigators. For example, Dammeyer, Nightingale, and McCoy (1997) found that 71% of clinicians expressed a strong belief that repressed memories exist and that prosecutors (48%) and defense attorneys (78%) were significantly less knowledgeable when it came to factors that would affect the accuracy of an eyewitness' testimony (Wise, Pawlenko, Safer, & Meyer, 2009). As noted, these (and other, see below) naïve beliefs stand in stark contrast to what the science of memory has shown about the manner in which memory operates (Arndt, 2012; Conway, Howe, & Knott, 2017; Howe, 2013a, 2013b, 2013c).

Despite this, many courts firmly assert the jury's ability to evaluate witness evidence (e.g., see New Jersey Courts [Eyewitness Instruction], 2012). In fact, research has shown that when evaluating the accuracy of witnesses' testimony, jury-eligible individuals rely on factors that are poor predictors of accuracy, ignoring factors that are good predictors (Pawlenko, Safer, Wise, & Holfeld, 2013). As memory-based testimony and identification of suspects are cornerstones of legal evidence, certain USA jurisdictions (e.g., in Arizona, New Jersey, Pennsylvania) have developed enhanced procedures (e.g., in the eyewitness identification of suspects) to help jurors weigh memory evidence. This is an important first step for the judicial system, considering that a single type of memory error, misidentification of suspects, is the largest single source of wrongful convictions in the USA. For instance, according to the Innocence Project, 2015,

> Eyewitness misidentification is the greatest contributing factor to wrongful convictions proven by DNA testing, playing a role in more than 70% of convictions overturned through DNA testing nationwide.
>
> *(retrieved from: www.innocenceproject.org/causes-wrongful-conviction/eyewitness-misidentification)*

In this chapter we examine the types of naïve beliefs that affect the judicial system and provide the scientific counterpoint to these beliefs. We begin with beliefs about memory in general, including a discussion concerning naïve beliefs about how memories are formed and subsequently retrieved (i.e., reproductive rather than reconstructive). We then turn to naïve beliefs about the effects of stress and trauma on encoding, storage, and retrieval, including views about repressed memories and their subsequent recovery. Finally, we discuss naïve beliefs about children's memories, how these memories are preserved into adulthood, and their subsequent reliability as evidence in the courtroom. Although we devote specific chapters to many of these topics later in this book, our focus here is on long-standing naïve beliefs about memory in these contexts.

Naïve beliefs about eyewitness memory

In a recent survey, Simons and Chabris (2011) asked 1500 members of the general public a series of questions about memory. Respondents were presented with statements such as, "Permanent Memory: *Once you have experienced an event and formed a memory of it, that memory does not change,"* and they had to indicate whether they agreed ("strongly" or "mostly") or disagreed ("strongly" or "mostly") with the statement, or indicate that they did not know the answer. These answers were compared with what the scientific consensus is concerning the correctness of each statement. Interestingly, for the majority of the statements (all but one) education level was negatively correlated with agreement. Table 3.1 shows the percentage agreement – "strongly" and "mostly" combined – for each statement by the general public and by the scientific experts. As can be seen in this Table, whereas the public mainly agreed with each of the statements, the experts mainly disagreed.

What these data clearly document is that people's beliefs about memory run counter to the scientific consensus about how memory

MYTHS AND NAÏVE BELIEFS ABOUT MEMORY

Excerpted from *Memories and Miscarriages of Justice*

actually operates. Because the general public serve on juries, these data make it clear that "common sense" notions about memory will not suffice when jurors must decide guilt or innocence based on memory evidence (for an early discussion of this general problem, see Deffenbacher & Loftus, 1982). However, such common-sense notions are frequently viewed as sufficient in courts of law. Indeed, as one judge (incorrectly) opined:

> Eyewitness testimony has no scientific or technical underpinnings which would be outside the common understanding of the jury; therefore, expert testimony is not necessary to help jurors "understand" the eyewitness' testimony.
>
> *(State v. Coley; 32 S.W.3d 831; Tenn. 2000)*

Consistent with this quote, studies have found that judges routinely overestimate a juror's ability to distinguish reliable from unreliable eyewitnesses as well as their common sense understanding of memory (e.g., Houston, Hope, Memon, & Read, 2013; Magnussen, Melinder, Stridbeck, & Raja, 2010).

Unfortunately, such overestimation is not uncommon even today. Indeed, the use of expert reports to guide the courts is still being shunned because it is believed that the scientific findings are "the same as, or very similar to, commonly held beliefs, common experience, and common sense" (Keane, 2010, p. 24).

MYTHS AND NAÏVE BELIEFS ABOUT MEMORY

TABLE 3.1 Percentage agreement to statements about memory by the general public (N = 1500) and memory experts (N = 89)

Statement	*General public (%)*	*Scientific experts (%)*
Amnesia. People suffering from amnesia typically cannot recall their own name or identity	82.7	0.0
Confident testimony. In my opinion, the testimony of one confident eyewitness should be enough evidence to convict a defendant of a crime	37.1	0.0
Video memory. Human memory works like a video camera, accurately recording the events we see and hear so that we can review and inspect them later	63.0	0.0
Hypnosis. Hypnosis is useful in helping witnesses accurately recall details of crimes.	55.4	0.0
Unexpected events. People generally notice when something unexpected enters their field of view, even when they're paying attention to something else	77.5	18.8
Permanent memory. Once you have experienced an event and formed a memory of it, that memory does not change	47.6	0.0

Data taken from Simons and Chabris (2011).

Related research has examined judges', jurors', and law enforcement personnel's understanding of issues related to eyewitness memory. Because there has been considerable research on eyewitness memory over the last 25 years, survey-based studies have been designed to assess the impact of this research on those involved in the legal system. Specifically, these studies have focused on ascertaining what prosecutors, judges, jurors, and the law enforcement officers know and believe about eyewitness memory (Wise et al., 2009). For example, Benton, Ross, Bradshaw, Thomas, and Bradshaw (2006) used a questionnaire originally designed by Kassin, Tubb, Hosch, and Memon (2001) to test eyewitness experts' knowledge of eye-witness memory. Benton et al. administered the same questionnaire to 42 criminal and civil state judges, 75 law enforcement officers attending training conferences, and 111 jurors drawn from individuals who had responded to a jury summons in Hamilton County, USA. The survey included 30 statements concerning eyewitness issues and three response options were provided for each statement ("Generally true," "Generally false," and "I don't know").

When comparing the responses with those of the eyewitness experts they found that for jurors, 87% differed in their responses. Jurors were the least knowledgeable on items relating to system variables (these are factors having to do with procedures used by the justice system, such as interviewing techniques or line- up procedures; see Chapter 7). They found that the officers' responses differed significantly from the experts' responses on 60% of the 30 eyewitness statements. These included wording of questions, effects of lineup instructions, presentation format (simultaneous or sequential), lineup fairness, and confidence malleability. Law enforcement officers and judges performed better but still not at the level of the eyewitness expert. Both disagreed with the experts on 60% of their responses. Disagreements were mainly focused on items related to the wording of questions, effects of lineup instructions, presentation format, description matching as opposed to suspect matching, and confidence malleability (Benton et al., 2006). Worryingly, law enforcement officers also differed significantly from the experts in relation to the effects of post-event information, child suggestibility, hypnotic suggestibility, child witness accuracy, false childhood memories, identification speed, and long-term repression. Benton et al. reported that the largest differences in law enforcement and experts' responses were found for child witness accuracy. The largest discrepancy between the responses of judges and experts was for hypnotic suggestibility, which was also found for juror responses.

It is important to note that such memory beliefs in these groups are not simply an aberration of sampling in a single study. In a comparison to Wise and Safer's (2004) survey of judges' knowledge of eyewitness testimony, on seven out of eight items that were similar in content, judges in both studies responded comparably. Items included confidence malleability (88% vs. 89%, respectively), weapon focus (67% vs. 69%), post-event information (81% vs. 84%), attitudes and expectations (86% vs. 94%), presentation format (29% vs. 19%), and the forgetting curve (41% vs. 31%) (Benton et al., 2006). These beliefs continue still. Even when law enforcement officers were from departments that have instituted eyewitness reforms (National Institute of Justice's Eyewitness Evidence: A Guide for Law Enforcement (hereafter "Guide"); Technical Working Group for Eyewitness Evidence, 1999), Wise et al. (2011) found that the reform and non-reform officers did not differ in either their knowledge of eyewitness factors or their use of proper interviewing procedures.

These findings show that there is still a large discrepancy between lay understanding of factors that will affect eyewitness accuracy and what we know from years of scientific research. This discrepancy is large and exists not only in jurors but also in judges, lawyers, and law enforcement professionals. Therefore, even those involved more directly with eyewitness evidence exhibit important limitations in their knowledge of eyewitness issues.

What is the solution? We have seen evidence that even with the existence of reforms and guides for law enforcement, these inaccurate beliefs still exist. However, more work is being done on the development and dissemination of standard protocols for

MYTHS AND NAÏVE BELIEFS ABOUT MEMORY

collecting evidence and also on instructions for jurors (Lampinen, Judges, Odegard, & Hamilton, 2005). However, as highlighted by Benton et al. (2006), the rest of the solution would appear to be the use of testimony from eyewitness experts where eyewitness evidence plays a pivotal role. It is clear that such knowledge needed to understand the risks associated with eyewitness memory accuracy is not common sense to everyone. As Benton et al. highlighted some 10 years ago:

> The legal system needs to become aware that the scientific and technical underpinnings of eyewitness memory research are not only outside the purview of common sense but also sufficient to warrant the admission of expert testimony as scientific knowledge. Consequently, eyewitness experts may be best able to serve and assist the court by providing information about the impact of system variables as delineated in the Guide, which is not common sense to jurors, judges or law enforcement personnel.
>
> (Benton et al., 2006, pp. 126–127)

However, who provides the expert testimony? In cases where the reliability of memory reports is an issue, psychologists are occasionally called as expert witnesses. Of course, it is not just the judiciary and laypeople that have naïve beliefs about memory, so too do some of those who provide clinical psychological advice in forensic cases. For example, Melinder and Magnussen (2015) surveyed 177 psychiatrists and psychologists using a similar questionnaire to that used by Wise and Safer (2004). Respondents had all served as expert witnesses in court. Their knowledge of eyewitness memory was compared with that of 819 psychiatrists and psychologists who had never served as expert witnesses. The

MYTHS AND NAÏVE BELIEFS ABOUT MEMORY

questions included beliefs about the link between eyewitness confidence and accuracy, the effects of post-event (mis)information, and whether memory for minor or peripheral details was an indicator of eyewitness accuracy. The results showed that many of the beliefs held by psychiatrists and psychologists were inaccurate, and those beliefs contradicted the consensus view of memory experts. Moreover, and contrary to their expectations, those who had testified about eyewitness memory in court did not perform any better on the memory questions than those who had not served as an expert witness. Worse, additional studies have shown that licensed clinical psychologists are no better than legal professionals and laypeople when it comes to understanding issues related to eyewitness memory (Magnussen & Melinder, 2012).

Further surveys conducted in the UK (e.g., Ost et al., 2013) and the USA (Patihis et al., 2014) have all shown similar patterns of naïve beliefs in licensed and unlicensed clinical psychologists. This survey research highlights significant false beliefs about memory that laypeople, clinical psychologists and psychiatrists, and those in the legal system (police officers, lawyers, and judge) labor under when eliciting memory evidence and when evaluating that evidence. Thus, the take home message of this research is that (clinical, counseling) psychologists are not memory experts just because they are psychologists. Professional psychologists in these samples, all of whom will have studied cognitive psychology to some extent at university, do not typically score higher than people in the legal field or those lay-people acting as jury members. Indeed, Magnussen and Melinder (2012) suggested that such results support the recommendations of the British Psychological Society Research Board's report *Guidelines on Memory and the Law* (2008), that memory expertise must be proved in each individual case and that courts are advised not to accept expert witnesses testifying on issues involving the reliability of memory reports unless expertise in memory research has been demonstrated.

Naïve beliefs about the effects of stress and trauma on memory

So far we have considered memory issues that relate to eyewitness accuracy. However, there are other memory issues surrounding the relationship between emotion and memory and the fate of traumatic memories where naïve beliefs can have serious consequences. One very prominent naïve belief is that when events are traumatic, memories for those events are particularly vivid, include considerable peripheral detail, and, depending on the extent of the trauma, are frequently repressed. Indeed, 81% of undergraduate students agreed with the statement that "traumatic memories are often repressed" (Patihis et al., 2014, p. 521). An additional 70% of these same students agreed that "repressed memories can be retrieved in therapy accurately" (Patihis et al., 2014, p. 521). Rubin and Bernstein (2007) also examined the beliefs of lay people for the plausibility that forgotten childhood sexual abuse among people could account for longstanding emotional problems and a need for psychotherapy. In a survey questionnaire, they asked respondents how likely they thought it was that a person with longstanding emotional problems and a need for psychotherapy could have been a victim of childhood sexual abuse. They found that only 17.8% said that this was implausible. Based on this finding, Rubin and Bernstein concluded that, "our results are important in their own right because they document a widespread belief in the general population of a Western Society that an event as memorable as childhood sexual abuse can be forgotten and still have marked effects on current behavior" (p. 777).

Magnussen and Melinder (2012) asked clinical psychologists to respond to the question, "Sometimes adults in psychotherapy remember traumatic events from early childhood, about which they previously had absolutely no recollection. Do you think such memories are real or false?" The results showed that 63% thought these memories were mostly real. Of course, it should be noted that this question does not require the respondent to agree or disagree as to whether the stories that abuse patients tell are generally true or false, but rather simply asks about beliefs regarding the truth of a recovered memory in therapy. Magnussen and Melinder argued that by agreeing with this statement it would seem to suggest a belief in the existence of a special repression or dissociation mechanism of forgetting.

Despite two decades of research on the veracity of recovered memories, it appears that beliefs today are similar to what they were some 20 years ago. Poole, Lindsay, Memon, and Bull (1995) surveyed clinical psychologists in the UK and the USA and found that 70% of the respondents have utilized various memory-recovery techniques with their clients to help uncover memories of early abuse. The evidence presented above seems to suggest that in many cases, lay people, and to some extent (clinical, counseling) psychologists, still believe in a special memory mechanism that leads to the repression of traumatic childhood memories.

Of course, such a belief runs contrary to the scientific consensus that has emerged in this area. Repression is not among the mechanisms of forgetting acknowledged by current memory science (McNally, 2003). In fact, well-controlled, large-sample studies have routinely showed that the victims of childhood sexual abuse remember these experiences into adulthood (Goodman et al., 2003) and that abuse severity is positively linked to memory longevity (Alexander et al., 2005). Thus, even very severe abuse tends to be remembered continuously throughout childhood and into adulthood. Although there may be some exceptions to this, the evidence suggests that such forgetting is likely due to normal mechanisms of failing to remember early childhood events (e.g., infantile and childhood amnesia) and, in cases of milder abuse (Alexander et al., 2005; Goodman et al., 2003), normal mechanisms of forgetting. As we showed in Chapter 2, there is no substantial or reliable evidence to support the existence of a special repressed memory mechanism (also see McNally, 2003).

Even more extreme beliefs about the effects of trauma on memory can be seen when we consider the notion of satanic ritual abuse (Bottoms, Shaver, & Goodman, 1996; Qin et al., 1998). The late 1980s and early 1990s saw a dramatic rise in the number of claims that people (particularly children) were subjected to satanic ritual abuse (see Chapter 2). Investigations into the reality of such claims in both the UK (La Fontaine, 1998) and the USA (Lanning, 1992) found no basis for the belief in such satanic ritual abuse. In spite of this, accounts of such abuse continued to appear (Scott, 2001). In fact, 32.4% of clinicians who were asked whether they had seen a case of satanic/ritualistic abuse said that they had (Ost et al., 2013). Unfortunately, it is the case that memories of satanic ritual abuse have almost always been reported as having been repressed and thus, many therapists have engaged in suggestive recovered memory techniques to "uncover" such claims. Despite the absence of empirical evidence to substantiate the existence of repressed memories of satanic ritual abuse or childhood sexual abuse, belief in their existence is remarkably widespread (see Lynn, Evans, Laurence, & Lilienfeld, 2015). For example, Poole et al. (1995)

found that 71% of clinical and counseling psychologists claimed to have encountered at least one case of repressed memory in their practice. Worse, what these beliefs can inevitably lead to is a very diverse group of therapeutic interventions designed to help the patient "accurately recover" these memories. Common to the various techniques (e.g., brainspotting, somatic transformation therapy, traumatic incident reduction; see Lynn et al., 2015) is that they tend to foster the development of false memories, not the recovery of true memories (see Chapter 2).

Naïve beliefs about children's memory

Survey research has also shown that when it comes to understanding children's ability to remember events, lay-people, clinical psychologists and psychiatrists, and those involved in legal services (e.g., judges, jurors, lawyers, police) share a number of false beliefs. For example, when asked "When a child's description of sexual abuse is disclosed over time, with more details being reported each time the child is interviewed, this indicates that the child's description is true," 52% of jury-eligible respondents agreed (Quas, Thompson, & Clarke-Stewart, 2005, p. 439). Disclosing more details across interviews is not indicative of children's accuracy when reporting abuse (see Chapter 8). Indeed, increasing detail with additional interviews is more an indication that suggestion is at work than that memory has "improved" (also see, Quas et al., 2005).

Curiously, 56% of respondents agreed that "a child cannot describe sexual abuse unless he/she actually experienced it" (Quas et al., 2005, p. 437). Of course, studies of children's knowledge of sexual matters shows that by age four, they have an understanding of some sexually relevant information (including the ability to name sexual body parts) and that sexually abused and non-abused children's rudimentary sexual knowledge does not differ (e.g., Gordon, Schroeder, & Abrams, 1990a, 1990b). Interestingly, only a minority of jury-eligible participants, 38%, agreed that children sometimes make up stories of having been sexually abused when they have not and 46% agreed that children can sometimes come to believe that they were abused when they were not actually abused (Quas et al., 2005).

It is clear that naïve beliefs about children's memory still exist for the majority of jury-eligible (and other) people. However, some progress has been made. For example, 71% of jury-eligible individuals agreed that "children are sometimes led by an adult into reporting that they have been sexually abused when they have not" and 70% agreed that "most children can be manipulated into making a false claim about sexual abuse" (Quas et al., 2005, p. 437). Fifty-eight percent of these same people also agreed that "repeatedly asking children specific questions, such as, 'Did he touch your private parts?' often leads them into making false claims of sexual abuse" (Quas et al., 2005, p. 437). Of course, whether it is a majority or a minority of individuals who hold naïve beliefs about children's memory, it is important to educate those jurors (judges, legal personnel) through expert testimony about the science of children's memory in order to ensure a just verdict (also see, Ceci & Friedman, 2000; Lyon, 2002).

Naïve beliefs about the reliability of adults' recollections of childhood events

Consider the following opinion rendered in a case of accusations concerning historic child sexual abuse:

MYTHS AND NAÏVE BELIEFS ABOUT MEMORY

Excerpted from *Memories and Miscarriages of Justice*

> It is difficult to see how . . . expert evidence can properly be tendered to establish a justifiable criticism of an adult witness who says that she suffered abuse throughout her childhood, which must have begun at too early an age for her to remember the first occasion [and who provided] highly specific details of abuse at such an early age . . . the jury should consider their own experiences, searching their recollections for their earliest memories, and analyzing what they could actually remember, and how far back their memories went. They did not require, and would not have been assisted by the evidence of an expert.
>
> *(R v. S; R v. W; 2006 EWCA Crim. 1404, Royal Courts of Justice, London, p. 9)*

Naïve beliefs about the ability of adults to remember early childhood events are rife. For example, consider one of the findings from Conway et al.'s (2014) survey of beliefs about memory in the general population. Respondents were asked to judge what was the age from which their own earliest memory dated and what was the age of other people's earliest memory. Figure 3.1 shows the distribution of first memories for one's own first memory (the open bars) and that of the estimated age of first memory of other people (the filled bars). What is interesting here is that the age of one's own first memory largely dates to the period of three to five years of age, in good agreement with many other studies of the age of first memories (e.g., Rubin, 2000). In contrast, however, the estimated age of the first memory of other people dates to the period from two to four years of age. In other words, people believe that other people generally have earlier memories than their own. This is a potentially serious overestimation. For instance, when a person is asked to judge the veracity of another person's earliest memories, as often happens in cases of historic childhood sexual abuse, then following (no doubt implicitly) the belief that other people have earlier memories than they do, they may have a bias to accept as possible, or even true, an account of a memory purportedly from a very early age.

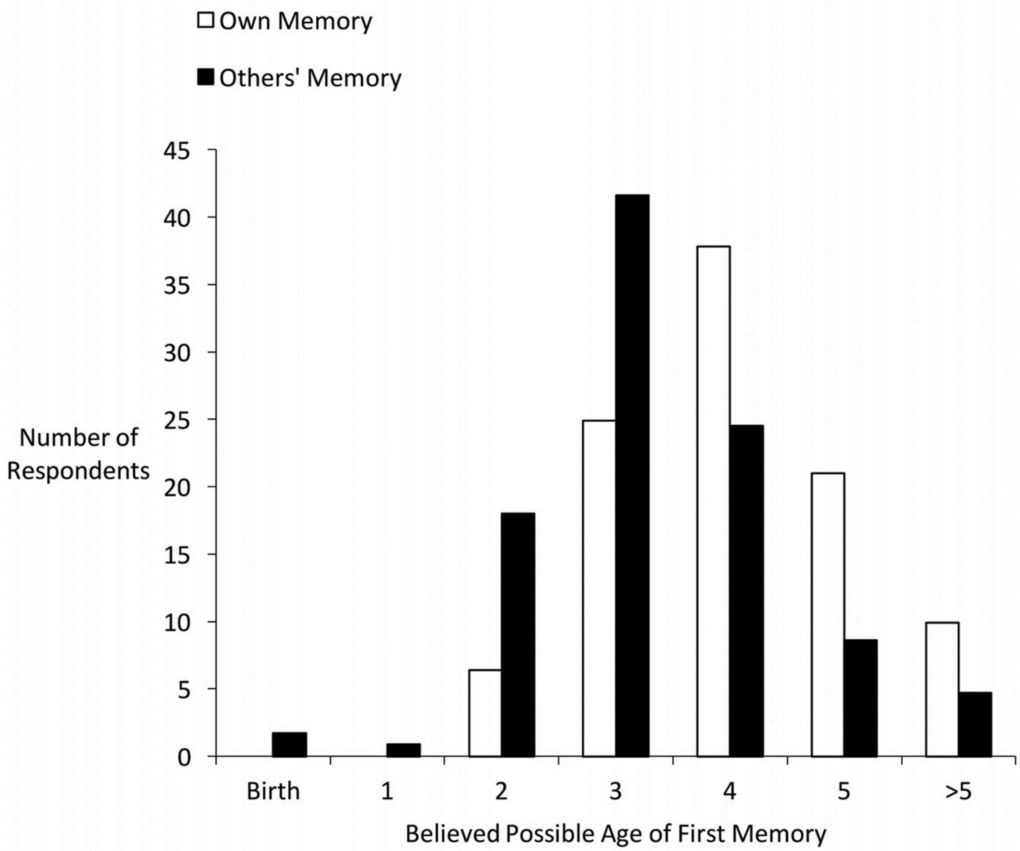
Nonetheless, although many respondents accurately state that remembering events from the first year of life is quite difficult, a significant minority of people, 15.1%, also believe that "with effort, we can remember events back to birth" (see Figure 3.1 and Patihis et al., 2014, p. 521). In a recent UK survey that compared memory beliefs of Chartered Clinical Psychologists, unchartered therapists, and first year undergraduate students in psychology it was found that there is still a strong belief that adults' reports of memories from a very young age are likely to be reliable and may have been unconsciously repressed (Ost et al., 2017).

We have already discussed naïve beliefs concerning trauma and repression, but a related question arises as to the reliability of early childhood memories that arise for the first time in therapy. A large majority of lay-people and professionals in Europe and the USA agree that such memories are likely to be true (see Melinder & Magnussen, 2015; Patihis et al., 2014). Of course, this belief too flies in the face of scientific evidence. Indeed, it is well known that traumatic events are often better remembered than more mundane events (e.g., McNally, 2003). Moreover, with the exception of events occurring during very early childhood (e.g., infantile amnesia) or events that are subject to alcohol- or drug-induced blackouts and documented brain anomalies (e.g., epilepsy), complete amnesia for traumatic events has been documented to be faked (see chapters in Christianson, 2007).

MYTHS AND NAÏVE BELIEFS ABOUT MEMORY

Excerpted from *Memories and Miscarriages of Justice*

FIGURE 3.1 Believed age of own and others' first memory.



Naïve beliefs about quality and quantity of evidence

One prominent naïve belief about memory accuracy is that accounts of events that are rich in detail are more accurate than accounts that are less detailed (Conway, Loveday, & Cole, 2016). When surveyed, laypersons (Magnussen et al., 2008), professionals from the legal field (Wise & Safer, 2004), and clinicians (Magnussen & Melinder, 2012; Melinder & Magnussen, 2015) appear to agree with this assumption. Here, participants were provided with the following statement; "A witness's ability to recall minor details about a crime is a good indicator of the accuracy of the witness's identification of the perpetrator of the crime" (Melinder & Magnussen, 2015, p. 56). The appropriate response, according to the most current science, would be, disagree. Here, it is well established in the scientific literature that false, not true, memories that can be frequently accompanied by such rich details (for a review, see Arndt, 2012). Results from the survey showed that only 33% of clinicians, 16% of laypersons, and 31% of judges (Melinder & Magnussen, 2015) made the correct choice to "disagree."

These findings provide further support for the need to educate both jury eligible laypeople as well as professionals involved in the legal system regarding the fragility of autobiographical memory and the well-documented factors that may undermine the reliability of human memory reports. It provides further evidence in support of the role of the expert witness who has a PhD and relevant expertise on issues involving the reliability of memory. This is an important point, because implausible memory performances are often misunderstood in the courts. Magnussen et al. (2010) review one recent example where the use of a memory expert may have prevented the wrongful convictions of Scottish citizens Thomas Campbell and Joseph Steele, who spent twenty years in prison for murder before finally being exonerated. The ultimate cause of the conviction was the impressive (but ultimately false) testimonies of four police officers who claimed to have overheard an allegedly self-incriminating remark by Campbell. The jurors and judge were so impressed with the confidence and detail of these testimonies, that they convicted Campbell and Steele of the murder.

To explain, in 1984, Thomas Campbell and Joseph Steele were sentenced to life imprisonment for armed assault and for murdering six people. In what is known as the Ice Cream Wars in Glasgow in the 1980s, there was a fierce competition for the most lucrative runs of ice cream vans with intimidation and violence employed by rival vendors. Campbell and Steele were convicted of shooting in the windows of the van owned by Andrew "Fat Boy" Doyle and for later setting his house on fire, resulting in the death of Fat Boy and five of his family members, including a baby. There was no evidence that the men had been at the site of the crime and the case rested on the testimonies of the four police officers who attended Campbell's home when executing a petition warrant that concerned the shooting, and who claimed to have overheard a remark by Campbell, "I only wanted the van windows shot up. The fire at Fat Boy's was only meant as a frightener which went too far." Campbell denied having made such a statement but the confident testimonies by the police officers obviously impressed the court.

However, it was precisely this point that worried the Scottish Criminal Case Review Commission. The commission noted that the police officers' accounts of the remark were identical, despite claims that they had not compared notes, and asked, what is the probability that four witnesses remember, in identical wording, a remark consisting of 23 words dropped under such conditions? The commission engaged a cognitive psychologist who conducted two experiments in which he tested the ability of witnesses to remember Campbell's alleged statement after being presented with a recording made with a Glasgow accent. None of the participants, including fourteen Scottish police officers, were able to remember the statement verbatim, and the majority of the participants remembered less than half the statement. The commission referred the case to the Scottish Appeal Court, the High Court of Judiciary, which decided that the evidence had been fabricated. In 2004, the court quashed the convictions of Campbell and Steele (cited in Magnussen et al., 2010, p. 1; taken originally from www.scotcourts.gov.uk/opinions/XC956.html). Magnussen and Melinder (2015) raise the question, would a memory expert have detected the unlikely memory performance of the police officers? They, like us, believe they would. Having such knowledge should help judges and jurors assess, with more skeptical eyes, the truly extraordinary memory feats that are sometimes presented in court.

Limitations of survey research

Much of the evidence discussed in this chapter concerning beliefs about memory has been extracted from studies using large-scale self-report surveys. There are, of course, several limitations to such surveys that should be highlighted before we conclude. First, as noted, these surveys reflect self-reports. Self-reports are not always a true reflection of actual belief or explicit behaviors. For example, results from Benton et al. (2006) may overestimate the extent to which police officers follow appropriate procedures for eyewitness interviews and eyewitness identifications.

Second, some of the statements could have been interpreted differently from how they were originally intended. If true, conclusions reached from responses may not accurately reflect what respondents actually believe. For example, take the statement used by Magnussen and colleagues (Melinder & Magnussen, 2015; Magnussen et al., 2012) regarding recovered memories in therapy; "Sometimes adults in psychotherapy remember traumatic events from early childhood, about which they previously had absolutely no recollection. Do you think such memories are real or false?" Responding "mostly real" leads us to conclude that psychologists believe in a special repressed or dissociative memory mechanism. Although this is a reasonable assumption to make (how else do respondents believe these memories can be true?), this proposition has not been explicitly put in this version of the question. Furthermore, such a question may also be too vague. Psychologists may be aware that recovered memory techniques can be suggestive and are susceptible to memory errors, but this question does not allow them to say this. Nevertheless, these survey results provide some very intriguing insights into people's (naïve) beliefs about memory.

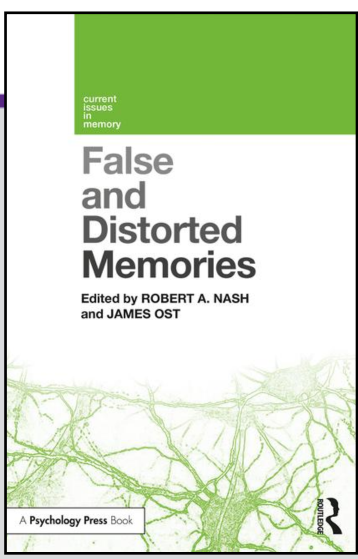
Conclusion

What lessons can we learn from this chapter? Evidence would suggest that, in general, knowledge about the functioning of human memory and its consequences in legal settings is quite poor. Lay-people, police officers, professionals from the legal field, and psychologists hold varying beliefs about critical issues of memory that are not supported by current memory science. When the reliability of memory is in question in the courts, this does have serious consequences (see earlier reference to the Innocence Project).

In the psychiatric and psychological field, mistaken beliefs regarding the complex workings of the human memory system are unlikely to lead to any serious consequences in everyday practice. However, when such false beliefs are brought into the courtroom, they may have catastrophic consequences (Melinder & Magnussen, 2015). As we have seen in Chapter 2, there have been a number of well-documented cases in both the USA and Europe where recovered memories of CSA have led to false convictions of innocent persons and, more so now, we are seeing cases of malpractice claims against therapists. We have also seen misguidance of the court from psychiatric and psychological expert statements relating to the reliability of witness memories when those reports are from individuals who are not memory experts (Goodman et al., 2007; Loftus & Ketcham, 1994). Thus, it is recommended that in order to prevent such wrongful convictions and reduce the number of false allegations caused by highly suggestive recovered memory (and related) techniques, it is important that psychiatric and psychological expert witnesses are updated on current memory science. They should also be aware of what constitutes as normal, memory performance for both children and adult witnesses, and also be able to show an understanding of the malleable nature of memory and the errors that can occur in memory performance. The famous Daubert ruling of the USA Supreme Court (*Daubert vs. Merrell Dow Pharmaceuticals, Inc.*) set the standards for governing expert testimony in court. It is recommended that "the opinions formulated by the expert witness are based on theories and methods that are testable, have been subject to peer review, have an acceptable error rate, and are generally accepted within the expert's community" (Melinder & Magnussen, 2015, p. 59). When it comes to remembering events, ones that play a critical role in convictions in the courtroom, there needs to be a clear set of guidelines concerning what constitutes a memory expert, ones that specify who should provide expert testimony in the courtroom. This, perhaps of all the lessons that can be learned from this chapter, is the most important one.

CHAPTER 3

COGNITIVE FLUENCY AND FALSE MEMORIES



This chapter is excerpted from

False and Distorted Memories

Edited by Robert A. Nash and James Ost.

Learn more

COGNITIVE FLUENCY AND FALSE MEMORIES

Excerpted from *False and Distorted Memories*

You are at a family dinner, in the middle of recalling a favorite childhood holiday, only to learn that you weren't actually there, or even born yet. In fact, what you were recalling as your own memory is actually someone else's. Besides embarrassment, you might also feel confused about why that content felt so easy to retrieve – after all, what you were remembering was a false not a real experience. From one moment to the next, people are aware of the ease or difficulty of ongoing cognitive activity – when we are trying to remember, when we are trying to imagine, or even when we are trying to read something, we have a sense of how easy the task feels (for a review see Jacoby, Kelley, & Dywan, 1989; Alter & Oppenheimer, 2009; Schwarz, 2010). And we use this subjective sense of processing, or cognitive fluency, to inform our judgements. Typically when something feels easy to process we think we have seen it before, we think it is trustworthy, frequent, safe, and true (e.g., Tversky & Kahneman, 1973; Begg, Anas, & Farinacci, 1992; Whittlesea, 1993; Song & Schwarz, 2009). And in the context of memory, information that is easily and rapidly recalled feels like it is the result of our own prior experience – for example, a family holiday – and thus part of our autobiog- raphy (Jacoby, Kelley et al., 1989).

But a feeling of fluency can mislead us: thoughts and images can spring to mind with the authenticity of prior experience, even when they are not real and not our own. In this chapter I examine the many routes to a feeling of fluent processing and the consequences for memory.

Cognitive fluency

Cognitive fluency is typically defined as the ease and speed with which people can process information (Jacoby, Kelley et al., 1989;

COGNITIVE FLUENCY AND FALSE MEMORIES

Alter & Oppenheimer, 2009). The major finding from this body of work is that people notice when things are processed quickly or feel easy, and they draw on this metacognitive experience as a source of information to guide their judgments, even when they have other available – more diagnostic – information to draw on. The storyteller at the dinner table could have drawn on factual details such as the timing of the event and his or her own date of birth to determine the authenticity of the memory, but instead the feeling of easy retrieval guided their reality check (see Schwarz et al., 1991, for evidence of people using retrieval ease over content to make decisions).

The various routes to a feeling of cognitive fluency

Of course, retrieval ease is just one route to a feeling of cognitive fluency. In fact, diverse manipulations can create a feeling of easy processing that people apply to inform their judgements of liking, safety, memory, truth, and even taste. Here I document some of the many routes to a feeling of fluent processing and the parallel effects that these manipulations exert on people's judgements.

Perceptual fluency

At the most fundamental level, the ease with which people can perceive stimuli can produce a feeling of cognitive fluency and influence their judgements. And a variety of manipulations can make information relatively easy to perceive. In a classic study, Reber and Schwarz (1999) showed subjects trivia claims against a white background presented in either high-contrast (e.g., dark blue) or low-contrast colors (e.g., yellow). When subjects evaluated the truth of the trivia claims, they were more likely to believe the claims that were easier to read – those presented in high-color contrast. Reading essays in an easy-to-perceive font can produce similar effects. Subjects thought students were more intelligent when their essays were presented in easy-to-read rather than *difficult-to-read* font (Oppenheimer, 2006). Even a brief prior exposure to an image or word can make that stimulus easier to perceive and feel more familiar (Whittlesea, Jacoby, & Girard, 1990; Whittlesea, 1993; Brown & Marsh, 2008). Although these manipulations are diverse, they all increase the ease of visually perceiving the stimuli and, as a consequence, produce a feeling of fluent processing.

Conceptual fluency

The ease of processing meaning can also bring about a feeling of cognitive fluency. Manipulations that promote rapid access to related concepts can make a target stimulus feel easy to understand (e.g., Kelley & Lindsay, 1993; Whittlesea, 1993). Seeing a related sentence (e.g., "The librarian reached for the top shelf and pulled down a book") before seeing a target word ("Read") can lead people to think the target word is more pleasant, perhaps because the preceding sentence activated related concepts making the target word more available in memory (Lee & Labroo, 2004). Similar effects can be seen with more complex stimuli; showing people an advert about mayonnaise can lead people to like a picture of a conceptually related product like ketchup (Lee & Labroo, 2004) and presenting trivia claims (e.g., "Macadamia nuts are in the same evolutionary family as peaches") along with a related photo (e.g., a bowl of macadamia nuts) can bias people to conclude that those trivia claims are true – even when the photo provides no diagnostic information about the claim (Newman, Garry, Bernstein, Kantner, & Lindsay, 2012; see also Labroo, Dhar, &

COGNITIVE FLUENCY AND FALSE MEMORIES

Excerpted from *False and Distorted Memories*

Schwarz, 2008).

Simple repetition is another source of conceptual fluency. We know from the truth effect literature that repeating claims can lead people to think they are true, and one component of this effect is that repetition can increase both perceptual (the claim is easier to identify perceptually when subjects have seen it before) and conceptual (the claim and related concepts are likely more accessible in memory after repetition) processing (Hasher, Goldstein, & Toppino, 1977; Bacon, 1979; Reber & Schwarz, 1999; Unkelbach, 2007; Dechene, Stahl, Hansen, & Wanke, 2010; Garcia-Marques, Silva, Reber, & Unkelbach, 2015).

Ease of imagery

The ease of mentally picturing a concept or experience is also tied to a feeling of fluency. For instance, in a classic study, ease of imagination influenced people's perceptions of the likelihood of a future event. Subjects thought it was more likely they would contract a disease when they were asked to imagine concrete (e.g., muscle aches, headaches) rather than abstract symptoms of the disease (e.g., dis- orientation, liver pain; Sherman, Cialdini, Schwartzman, & Reynolds, 1985). The ease with which people can construct a mental image also influences their future intentions. In one study, subjects said they were more likely to purchase an advertised trip if the brochure made it easy to imagine the destination (subjects saw either a clear and high-quality or a partially blurry and low-quality image of the destination; Petrova & Cialdini, 2005). This particular instantiation of cognitive fluency likely incorporates both perceptual and conceptual processing ease.

COGNITIVE FLUENCY AND FALSE MEMORIES

Excerpted from *False and Distorted Memories*

Linguistic ease

The ease of processing linguistic information can also give rise to a feeling of fluent or disfluent processing. The ease or difficulty of pronunciation can have reliable effects on people's assessment of chemicals, stocks, amusement park rides and even people. When names are easy to pronounce (e.g., Ohanzee, as opposed to Tsiischili), people conclude that the bearer of the name (whether it be a chemical, stock or person) is safe, familiar and likely a credible source of information (Alter & Oppenheimer, 2006; Song & Schwarz, 2009; Laham, Koval, & Alter, 2012; Newman et al., 2014).

In most cases, the interpretation of cognitive fluency is not dependent on the way it came about (see Alter & Oppenheimer, 2009, or Schwarz, 2015 for a review; cf. Lanska, Olds, & Westerman, 2014). That is, fluency can arise from a perceptual, conceptual or even a linguistic source and it can have similar effects on people's judgements. But how people make sense of cognitive fluency is flexible, and can depend on the domain of judgement, the features of the task at hand and expectations and biases an individual carries with him or her (Jacoby, Kelley et al., 1989; Schwarz, 2004; Alter & Oppenheimer, 2009; Whittlesea, 2011).

Interpretation of cognitive fluency

In the domains of belief and memory, fluency tends to signal authenticity. Indeed, a large literature demonstrates that when something feels quick and easy to process, people tend to conclude that the thing is familiar or that they have seen it before (e.g., Jacoby & Dallas, 1981; Whittlesea, 1993; Lanska et al., 2014; Brown & Marsh, 2008). Using a feeling of fluency in this way makes sense because things that we have seen before usually are easily

COGNITIVE FLUENCY AND FALSE MEMORIES

Excerpted from *False and Distorted Memories*

and more quickly identified the next time we encounter them. Related literature on belief tells us that when something feels quick and easy to process, people tend to conclude that the thing is true (Hasher et al., 1977; Bacon, 1979; Reber & Schwarz, 1999; Unkelbach, 2007; Hansen & Wänke, 2010). This also makes sense. People tend to make truth judgements based on (their perceptions of) social consensus – following the rule that if lots of people believe it then it is probably correct (see Schwarz, 2015). And there is a relationship between social consensus and fluency; when lots of people believe something, you have probably heard it many times and it should feel quick and easy to process or bring to mind (Weaver, Garcia, Schwarz, & Miller, 2007; Schwarz, 2015). In fact, a feeling of fluency should facilitate a "yes" response to several questions we tend to ask ourselves when we assess truth – whether something is credible, coherent and consistent and whether there is supporting evidence should all benefit from a feeling of easy processing (see Schwarz, 2015 for a review).

So when things feel fluent, people usually conclude they are familiar and true. But a large body of work also suggests that a feeling of fluency is experienced as inherently positive, so when things feel fluent, people usually conclude they are good (Winkielman, Schwarz, Fazendeiro, & Reber, 2003; Reber, Schwarz, & Winkielman, 2004). There are several lines of evidence for this idea. First, people's responses fit with the idea that fluency increases positive but not negative evaluations – a feeling of fluency increases judgements of liking and beauty, but does not boost ratings of disliking or ugliness (Reber, Winkielman, & Schwarz, 1998; Seamon, McKenna, & Binder, 1998). Second, a feeling of fluency can lead people to smile, a physiological marker of positive affect (Winkielman & Cacioppo, 2001; Topolinski, Likowski, Weyers, & Strack, 2009). Third, more recent research has shown that a feeling of fluency can influence sensory perceptions along positive dimensions – leading people to agree more with a claim that a wine is high quality and, after sampling a wine, conclude that it actually tastes better too (Cardwell, Newman, Garry, Mantonakis, & Beckett, in prep).

Taken together, this work suggests that fluency is a uniquely positive cue, but there is also evidence that fluency can be interpreted broadly depending on expectations and beliefs of a person experiencing fluency, as well as the contextual features of a situation (see naive theories, Schwarz, 2004; see also Jacoby, Kelley et al., 1989; Unkelbach, 2006). For instance, the fluency–truth link can be reversed if people learn, in a controlled experiment, that fluency is indicative of falseness rather than truth (Unkelbach, 2007). We also know that the context of a judgement can influence how people interpret an experience of fluency – a wine label with difficult-to-process font might seem interesting, expensive and appealing, whereas an essay written in difficult-to-process font might seem poorly written by an uninspired author (e.g., Oppenheimer, 2006; Mantonakis, Galiffi, Aysan, & Beckett, 2013). The fact that fluency comes from a variety of sources and can be interpreted flexibly makes it a powerful metacognitive cue. What's more, because we draw on fluency to evaluate not only whether something is familiar, but also whether something meets the criteria of a real memory, it has the power to shape how we remember our recent and distant pasts.

Fluency and consequences for memory

We know from our own experience that memory is not a perfect record of the past. The story at the dinner table tells us that memory is a mix of true information, details one heard from

COGNITIVE FLUENCY AND FALSE MEMORIES

Excerpted from *False and Distorted Memories*

someone else, saw in a photo or even something one imagined. And because memory has this reconstructive quality, we have to sort it, deciding whether something is the result of real experience or whether it is the result of a suggestion, our own imagination or another external source (see Garry & Hayne, 2013, for a review). We sort through memory using a process called source monitoring whereby we evaluate whether information that has come to mind looks and feels like a real memory (Johnson, Hashtroudi, & Lindsay, 1993; Lindsay, 2008; Lindsay, 2014). We review qualities such as sensory characteristics and contextual information (asking whether it is vivid and easily imagined) and conceptual coherence (asking whether it fits with other events that happened and whether it is plausible). And if a mental event meets one's criteria on these variables, one would conclude it is the result of prior experience – a genuine memory (Johnson et al., 1993; Lindsay, 2008; Lindsay, 2014).

Ease of processing bears on each of these evaluations. When a mental event is rich in sensory characteristics, it should be easy to mentally picture; when an event has a lot of contextual detail, it should be easy to access the meaning; and when an event is coherent with one's life story or schema for a particular situation, it should feel conceptually fluent (see also Jacoby, Kelley et al., 1989). That is, a feeling of fluency along any of these dimensions should support conclusions about memory authenticity. This fluency–reality link should serve people well most of the time. But there are two key reasons why relying on feelings of fluency might make people vulnerable to making mistakes in memory and in some instances vulnerable to developing wholly false memories.

1 Fluency can arise without prior experience (via conceptual, perceptual or even linguistic surface features of a stimulus or experience) and so create a sense of familiarity, making us feel like we have seen something before when we have not. For instance, in one study people saw target words embedded in semantically related sentences (e.g., "The anxious student wrote a *test*") or semantically unrelated sentences (e.g., "Later in the afternoon she took a *test*"; Whittlesea, 1993). When people saw words in related contexts – when the words felt conceptually fluent – subjects were more likely to say they had seen the word before, even when they had not. In this instance, the source of fluency was simply the conceptual context in which the word appeared, not previous exposure. In another study, subjects examined a series of news headlines and were asked to rate whether they remembered hearing about each news event (Strange, Garry, Bernstein, & Lindsay, 2011). When the headlines were paired with a related photograph, people were more likely to say they remembered the news event, even when the event never truly happened in the first place (e.g., a false event that "[Tony] Blair under fire for botched Baghdad rescue attempt; won't step down."). This tendency to feel like they remembered the news event was likely driven by the conceptual ease of picturing and thinking about the event when the headline appeared with a related photo. Again, the source of fluency was not prior experience, rather it was the context in which the target stimulus was evaluated.

The consequences of fluency can reach beyond simply thinking we have seen something before. A feeling of fluency that is produced in the moment can also lead people to think they have experienced something in the past. In one study subjects saw adverts for popcorn; when the adverts made it especially easy to imagine the product (using relatively more concrete language), subjects were more likely to (falsely) believe they had experienced the product –

COGNITIVE FLUENCY AND FALSE MEMORIES

Excerpted from *False and Distorted Memories*

a false experience effect (Rajagopal & Montgomery, 2011). In each of these studies, the source of fluency was not prior experience, rather it was produced via attributes of the stimulus that made it easy or difficult to complete the task at hand – thinking about a word, headline or product. Unless attention is drawn to the source of processing fluency, people are inclined to use a fluency experience as evidence for the judgment at hand – in these instances, evaluations of familiarity and prior experience (see also following section on discounting).

2 **We are often bad at tracking the source of fluent processing**. Even if we have seen something before, we may incorrectly interpret the reason or source of easy processing. In a recent study, subjects judged the familiarity of a series of faces. The key manipulation was that the faces appeared with a novel, familiar, or a neutral background. When subjects saw faces in front of a known background (e.g., a famous landmark) they were more likely to say they had seen the accompanying face before (Deffler, Brown, & Marsh, 2015). That is, people were bad at tracking the source of the fluent processing. Although they would be correct in concluding that they had seen the landmark before, the fluency of the landmark leaked onto their judgements about the familiarity of the face. Another example of this kind of error is in a study in which people were shown a series of health-related myths and facts. In an initial phase, when subjects saw each claim, they were told whether the claim was true or false (Skurnik,Yoon, Park, & Schwarz, 2005). When people were tested immediately for their memory of the claims, they were good at recalling which claims were true and which were false. But when people were tested after a delay (when it was more difficult for people to recall all of the details), they remembered the myths as being true – and people were especially likely to make this mistake when they had repeatedly learned that

COGNITIVE FLUENCY AND FALSE MEMORIES

a claim was false. In other words, seeing the claims earlier made them later feel fluent and familiar, and people did not track the reason for that fluency – which was that they had repeatedly learned that the statements were myths. Garry and colleagues provide a more dramatic example of how losing the source of fluency can lead to false beliefs. In their study, subjects first rated whether they had experienced a series of life events. Then subjects imagined some events, but not others (Garry, Manning, Loftus, & Sherman, 1996). When subjects were tested again in phase 2, they were more likely to believe they had experienced events that they had really only imagined – perhaps because those events now contained more contextual information and were easier to mentally picture. In this instance subjects have failed to identify the reason for fluent processing – that imagining and elaborating on the event likely made it more available in memory, easy to retrieve and relatively more detailed than events they did not imagine.

The studies reviewed thus far have shown that fluency can lead people to mis-remember words and claims, increase their belief in childhood events and remember using products they have not encountered. But fluency can contribute to more dramatic departures from reality. In a classic series of studies Loftus and colleagues demonstrated how people can come to remember whole events that never actually occurred (Hyman, Husband, & Billings, 1995; Loftus & Pickrell, 1995; Mazzoni, Loftus, & Kirsch, 2001). In these studies, subjects usually read four childhood events, three that were true and were provided by the parents and one that was false and never actually happened. The descriptions of the false events looked just like the real ones, containing familiar people and places, for instance: "You, your mom, Tien, and Tuan all went to the Bremerton K-Mart. You must have been 5 years

COGNITIVE FLUENCY AND FALSE MEMORIES

Excerpted from *False and Distorted Memories*

old at the time . . . and somehow lost your way in the store. Tien found you crying to an elderly Chinese woman . . ." (Loftus & Pickrell, 1995, p. 721). In an attempt to recall the events, subjects read the event descriptions and tried to imagine the events; in some studies subjects also reviewed photographs to help "cue" their memories (e.g., Wade, Garry, Read, & Lindsay, 2002; Lindsay, Hagen, Read, Wade, & Garry, 2004). After repeated interviews an average of 20 per cent (across 12 studies; see Newman & Garry, 2014) of subjects came to remember the false event as though it was a real memory. Although in these studies many factors are operating to induce false memories, many of the techniques – repetition, photographs, guided imagination and familiar event details – likely enhanced processing ease and facilitated errors in source monitoring.

Fluency can also have consequences for how we perceive our memory system. In one study, subjects recalled either 4 or 12 memories from their childhood (Belli, Winkielman, Read, Schwarz, & Lynn, 1998). When they were subsequently asked to evaluate their memory system and judge whether there were gaps in their memory, those people who had recalled 12 memories thought they had more holes in memory, even though they actually recalled more. This finding has been interpreted using a cognitive fluency account – that is, people used the ease of retrieval to judge their memories, rather than the quantity or content of what they recalled (Belli et al., 1998). When we can retrieve just a few instances easily, we feel like we have more memories than when we recall many memories with effort. In this situation, fluency did not produce false memories; it simply swayed people's estimates about the completeness of their memory. But concerns about the completeness of one's memory may leave people especially vulnerable to suggestion or attempts to fill in the perceived gaps.

COGNITIVE FLUENCY AND FALSE MEMORIES

Excerpted from *False and Distorted Memories*

Discounting experiences of easy processing

One way people can protect themselves from the effects of easy processing is to find out where the effects came from. The influence of fluency is dramatically decreased and even reverses when people can identify the source of fluent or disfluent processing. For instance, perceptual fluency effects disappear when people can identify the reason for visual difficulty (Oppenheimer, 2006). In one study, subjects evaluated an applicant's personal statement for admission into graduate school. Half of the subjects received a low-toner version of the personal statement – it was difficult to read, clearly because the printer had almost run out of ink. The other half received a high-toner version of the personal statement. Based on a fluency account, subjects should have been less willing to accept the low-toner applicant. But the reverse occurred; people were more likely to recommend the low-toner applicant than the high-toner applicant. The explanation for this effect is that people easily identified the source of processing fluency and responded in a way to compensate for its potential influence on their judgement. A similar effect occurs with ease of retrieval. In a different study, Oppenheimer (2004) asked subjects to evaluate the popularity of a series of surnames. When subjects came across the surname "Bush" they tended to underestimate its popularity. Given that Bush should be a relatively familiar last name (particularly for Americans), it is surprising that subjects did not assign a higher frequency estimate (cf. Jacoby, Woloshyn, & Kelley, 1989). Oppenheimer explained his results by suggesting that people identified the source of processing fluency and overcompensated for its possible influence. That is, people thought that "Bush" felt familiar because of President George Bush, and so assumed that fame was biasing their assessment of frequency. That is, they identified the source of fluency and discounted it in their judgement. There is also evidence of more fine-grained detection of the source of fluent processing. In a classic study by Jacoby and Whitehouse (1989), subjects saw a series of words in an initial study phase. In a subsequent test phase, they saw a mix of old and new words and had to decide whether they had studied them earlier. The key manipulation was that during the test phase, some of the target words appeared after a short exposure to a matching or nonmatching prime word. When the prime and the target words matched, people were more likely to (falsely) claim they studied the target word in phase one – perhaps because of an increase in perceptual fluency. But this effect disappeared when subjects were aware of the prime. That is, when the prime word appeared for 200 ms and people were aware of its presence, they identified the source of fluency and no longer used the fluency provided by the prime word in their judgement.

In each of the examples noted here, the subject has identified the source of processing fluency and attempted to combat it. But the same discounting is also possible if people are explicitly alerted to an alternative explanation for their fluency experience (Schwarz et al., 1991; Novemsky, Dhar, Schwarz, & Simonson, 2007). For example, subjects will no longer use retrieval ease to make judgments about memory if they think the retrieval experience (easy or difficult) is due to music that the experimenters played in the background whilst subjects attempted to recall memories (Schwarz et al., 1991).

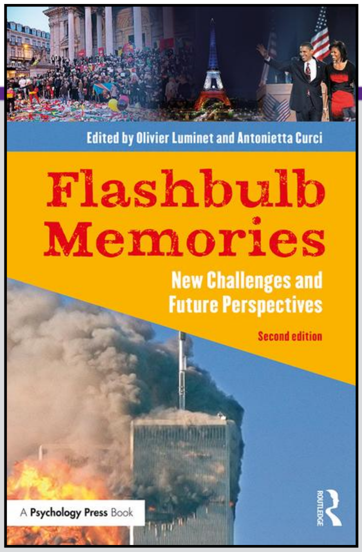
Concluding remarks

A feeling of easy processing can be a useful cue to identify real experience. Indeed, when something actually happened, fluency (regardless of how it is elicited) is diagnostic of memory. But when an event did not occur and fluency arose via characteristics of a stimulus that influenced the visual, conceptual and linguistic ease of processing, it can lead us astray. How can people protect themselves from these effects? One way is to have a critical lens at retrieval (see Johnson et al., 1993; Lindsay, 2008, 2014). Indeed the influence of fluency is reduced when people have the available cognitive resources to reason and critically examine what is real and what is not (e.g., Jacoby, Woloshyn et al., 1989). A feeling of (dis)fluency might help too. There is growing evidence that a feeling of difficulty can actually trigger a more analytical processing style. When people experience disfluency it can inform them about a judgement target – that it is false or untrustworthy (e.g., Newman et al., 2014), but an experience of difficulty can also change how people approach a task, facilitating error detection and source monitoring (Song & Schwarz, 2008; see also Hernandez & Preston, 2013). In fact, people can produce a feeling of disfluency themselves – simply furrowing one's eyebrows can make a cognitive task feel difficult (e.g., Strack & Neumann, 2000). Perhaps we could use this kind of strategy to trick ourselves into more careful analysis of mental events, reducing our susceptibility to false memories – and protecting ourselves from embarrassing moments at family dinners.

CHAPTER

4

ORDINARY MEMORY PROCESSES SHAPE FLASHBULB MEMORIES OF EXTRAORDINARY EVENTS: A REVIEW OF 40 YEARS OF RESEARCH



This chapter is excerpted from

Flashbulb Memories, 2nd Edition

Edited by Olivier Luminet and Antonietta Curci.

Learn more

ORDINARY MEMORY PROCESSES SHAPE FLASHBULB MEMORIES OF EXTRAORDINARY EVENTS

A REVIEW OF 40 YEARS OF RESEARCH

Excerpted from *Enter Book Title*

We review the shifting definition of "Flashbulb memory" in the 40 years since Brown and Kulik coined the term. We evaluate evidence for veridical, phenomenological,and metacognitive features that have been proposed to differentiate Flashbulb from ordinary autobiographical memories. We further consider the event conditions thought to be necessary to produce Flashbulb memories and discuss how post-event processing may distinguish Flashbulb memories. We conclude that a categorical dissociation between flashbulb and other autobiographical memories is untenable, but that Flashbulb memories still pose important, as yet unanswered, questions.

Brown and Kulik (1977) observed a phenomenon that captured the public's attention – seemingly indelible memory for important, emotional events. They dubbed it "Flashbulb memory (FBM)" and conducted the first modern empirical study on the topic. The concept was equally effective in capturing the attention of memory researchers, and in the 40 years following their seminal publication, the topic has been investigated almost as often as the events that lead to such memories allow. During this time, the description of the phenomenon has undergone an interesting and important transformation.

The initial "special mechanism" hypothesis was that FBMs were a permanent, veridical (though not necessarily complete) memory record that resulted from a unique memorial process involving automatic encoding of all aspects of an important (emotional) event as it happened. However, this strong hypothesis did not last long, as evidence of both errors of omission and commission in the recall of FBMs were soon identified (Christianson, 1989; Neisser & Harsch, 1992; Neisser *et al.*, 1996). Consequently, the revised FBM hypothesis was more agnostic as to why they developed and how they were different from other autobiographical memories (AMs).

ORDINARY MEMORY PROCESSES SHAPE FLASHBULB MEMORIES OF EXTRAORDINARY EVENTS

A REVIEW OF 40 YEARS OF RESEARCH

Excerpted from *Flashbulb Memories*

Our task is to determine if FBMs are simply a subtype of autobiographical remembering or if they are a distinct category of memories. If there are special mechanisms involved in FBMs then there should be three factors that differentiate them from everyday memories. First, with regards to memory properties, FBMs should be different from ordinary AMs in some way; they could be more accurate or more vivid, for example, than everyday memories. The second way to identify FBMs is that the conditions necessary to produce these memories should be different from ordinary events; for example, they could require strong emotions. Finally, how the individual processes the event (e.g., how one rehearses the memory) should differ for FBMs relative to AMs. For each of these, the claim of a special mechanism requires more than just a difference that could be seen as one extreme of a continuum; there should be some discontinuity between "ordinary" memories and "special" FBMs. In other words, there should be some *a priori*, objective threshold for defining whether a given memory meets the criteria for being considered an FBM.

In order to compare FBMs to ordinary AMs we need to define "ordinary." Here, we will define ordinary AMs as easy-to-access memories that are brought to mind by a request for a particular kind of memory (e.g., a memory from a particular time, of a particular type of event, or in response to a particular word). The results would likely be different if FBMs were compared to trivial or noteworthy memories, but "trivial" and "noteworthy" require defining along which dimension the events are trivial or noteworthy. Unless noted otherwise, the comparisons we report from the literature are between ordinary AMs (as defined above) and FBMs.

We will proceed to review the various ways in which FBMs have been claimed to be different from AMs and the various mechanisms proposed to cause these differences. We will ask if FBMs have more of a given property and report whether there are consistent findings across studies showing that FBMs are more extreme. This is the minimal test of a special mechanism. If this test is met, we will ask if the differences are large enough to exclude a continuum on which FBMs are at one end and where there is little overlap in the distributions. For the proposed mechanisms, we will also ask whether they have been shown to be necessary (i.e., can FBMs exist when these mechanisms are not invoked), and whether they have been shown to be sufficient (i.e., can FBMs occur only when these mechanisms occur). As a summary, our conclusions are indicated in Tables 4.1 and 4.2. Table 4.1 has a list of the ways FBMs have been claimed to be unique. Table 4.2 describes the mechanisms proposed to account for these differences.

TABLE 4.1 A summary of the differences between FBM and ordinary AM

	FBM > AM		
	No	*Yes*	*Discontinuous*
Memory characteristics			
Longevity	X		No
Accuracy	X		No
Consistency	X		No
Vividness		X	No
Confidence		X	No

Note: Discontinuous implies a large difference with little overlap in the distributions; no characteristics exhibit this. We would have the same results if we replaced "discontinuous" with "as compared to noteworthy memories that were not in response to a flashbulb event."

TABLE 4.2 A summary of the evidence supporting the mechanisms proposed for enhancing FBM relative to ordinary AM

	Sufficient	*Necessary*	*Only in FBM*
Event conditions			
Consequentiality	No	No	No
Distinctiveness	Yes	No	No
Negative affect	No	No	No
Memory processes			
Significance	Yes	Yes	No
Surprise	No	No	No
Emotional intensity	?	No	No
Rehearsal	?	?	No

Characteristics of FBMs

Longevity

Surprisingly, there is a paucity of evidence in the FBM literature addressing the relative permanence of such memories. Anecdotally, FBMs are extremely long lasting. Empirically, studies that examine longevity typically obtain one retrospective report years after the event and evaluate it for vividness and completeness to determine whether it qualifies as an FBM. For example, using these criteria, between 50% and 99% of participants in any given study have FBMs for learning of historical events after delays of approximately ten years (Brown & Kulik, 1977; Hirst *et al.*, 2015), approximately 30 years (Colegrove, 1899), and approximately 60 years (Berntsen & Thomsen, 2005).

Berntsen and Thomsen (2005) asked participants to recall another event from the same time period (i.e., the most positive and most negative personal event during the occupation period) and found that most were able to do so. Contrary to the claims of Brown and Kulik (1977), there are other events that one can remember from equally long ago. Denver, Lane, and Cherry (2010) provide converging evidence for the objective longevity of FBMs without a relative superiority to ordinary AMs. When their participants were provided with a description of the FBM phenomenon and asked to freely recall public events for which they have this type of memory, their recall produced a standard reminiscence bump (Rubin & Schulkind, 1997). Denver, Lane, and Cherry (2010) further showed few reliable differences between older adults' and younger adults' FBM despite vast differences in the delay intervals among events recalled by both groups. In conclusion, FBMs are long-lasting, but they are not indelible, nor are they are more permanent than noteworthy everyday memories.

Accuracy

Talarico and Rubin (2009) describe in detail the (surprisingly) few studies that investigate verifiable accuracy of FBMs. The two case studies report at least one critical inaccuracy, though many accurate details (Greenberg, 2004; Neisser, 1982; Thompson & Cowan, 1986). Furthermore, those inaccuracies are systematically biased to enhance the features of the event that contribute to its personal significance. The day of Danish liberation is remembered as more sunny, less cloudy, less windy, less rainy, and/or warmer than it actually was by Berntsen and Thomsen's (2005) participants, and Neisser (1982, 1986) misidentified a football game as a baseball game (i.e., America's pastime) when remembering the attack on Pearl Harbor (Thompson & Cowan, 1986). The issue of objective accuracy is ripe for further investigation. Event features most likely to be recalled accurately, as well as the magnitude and direction of errors, should be examined whenever archival data are available to confirm self-reports (see Luminet & Spijkerman, in press, for a similar argument). A larger body of evidence has examined consistency between memory reports as a proxy for accuracy.

Consistency

In order to obtain consistency data, two (or more) retrospective reports are collected. The report closer in time to the event is considered the standard. Later reports are then compared to that earlier report and inconsistencies are identified. Although two consistent reports are not necessarily accurate, an inconsistent report implies that at least one report is inaccurate. FBM consistency has also been shown to correlate with performance on the autobiographical Implicit Association Test (aIAT), itself a correlate of accuracy (Curci *et al.*, 2014; Lanciano *et al.*, 2013). Contrary to the arguments of Julian, Bohannon, and Aue (2009), we do not consider "wrong time slices" (Neisser & Harsch, 1992) to be accurate. Recalling an event that actually occurred (e.g., a 30th birthday party) but was not the event requested (e.g.,"tell me about your 40th birthday party") is inaccurate recall. Because changes across recalls do not imply a lack of accuracy, investigators may want to emphasize explicit inconsistencies when drawing conclusions about FBMs.

The overwhelming evidence is that FBMs are incomplete (Brown & Kulik, 1977) and include inconsistencies (Christianson & Engelberg, 1999; Curci, 2005; Curci & Luminet, 2006; Curci *et al.*, 2001; Greenberg, 2004; Larsen, 1992; Lee & Brown, 2003; Liu,Ying, & Luo, 2012; McCloskey, Wible, & Cohen, 1988; Nachson & Zelig, 2003; Neisser, 1982; Neisser & Harsch, 1992; Schmolck, Buffalo, & Squire, 2000; Talarico & Rubin, 2003, 2006; Weaver, 1993; Weaver & Krug, 2004; Wright, 1993). Memory for the general gist of the event is better than memory for specifi details (Bohannon & Symons, 1992;

Pillemer, 1984; Schmidt, 2004, 2012; Schmidt & Bohannon, 1988; Thompson & Cowan, 1986) and peripheral details are more likely to be inconsistent than are central details (Christianson, 1989; Romeu, 2006;Tekcan *et al.*, 2003). However, how one divides responses into central vs. peripheral information often coincides with whether the memory reports show consistency or inconsistency. Tekcan *et al.* (2003) considered "time" and "others present" to be peripheral details because those two questions were responsible for the majority of inconsistencies found in their participants' memory reports. Which categories account for reliable recall is also variable. For example, both Christianson (1989) and Pillemer (1984) found reliable recall for "informant," and "location," information. Pillemer (1984), but not Christianson (1989), found "ongoing activity" to be reliably recalled. Importantly, in none of these investigations was recall of central vs. peripheral details of a non-FBM obtained. In Curci and Luminet's (2009) study of French President Mitterrand's death, French participants were more consistent in location and time details than were Belgian participants, even though both groups showed high overall consistency within these categories.

Moreover, inconsistencies in FBMs, once introduced, are repeated over time and not corrected or further altered (as they are for semantic details of the event) (Coluccia, Bianco, & Brandimonte, 2006; Hirst *et al.*, 2015; Tekcan *et al.*, 2003; Weaver & Krug, 2004). Further, delayed recall is highly related to initial recall (Conway *et al.*, 2009; Weaver *et al.*, 2008). Longer delays between the event and the initial memory report often produce enhanced consistency scores (Coluccia, Bianco, & Brandimonte, 2006; Weaver *et al.*, 2008; Winningham, Hyman, & Dinnel, 2000), though not always (Coluccia, Bianco, & Brandimonte, 2010; Kvavilashvili *et al.*, 2009; Lee & Brown, 2003). Within AMs, the total number of consistent details has also been shown to increase with the number of rehearsals (Campbell *et al.*, 2011; Nadel, Campbell, & Ryan, 2007). Lastly, FBMs do not include fewer inconsistencies than everyday memories (Talarico & Rubin, 2003, 2006).

Therefore, in conjunction with the evidence described above, we must conclude that FBMs are not permanent, perfect copies of experienced events, nor are they so much more consistent than everyday memories that a special mechanism is required to differentiate the categories.

Vividness

However, there are characteristics of FBMs that may still differentiate them from ordinary AMs. Vividness has been of interest to FBM research since Brown and Kulik (1977) described the "live quality that is almost perceptual" (p. 74). Julian *et al.* (2009) have suggested that elaboration (i.e., quantity of details recalled) may be of interest as well1. There are strong pragmatic advantages to assessing elaboration as it only requires obtaining a single memory report, irrespective of delay from the event. However, there are at least three methodological issues that undermine its utility. First, there is wide variability in both the total number of details and the nature of those details when defining FBM. Kizilöz and Tekcan (2013) asked a large sample of individuals to recall three separate FBMs and identified, without regard to accuracy or consistency, seven distinct categories that captured the majority of information provided in the narratives: *informant*, *location*, *others present*, *own ongoing activity*, *own affect*, *own thoughts*, and *aftermath*. Interestingly, although *time* had been asked of participants in more than half of the studies included in their review (as were *location*, *ongoing activity*, *source*, and *others* *present*), it was rarely spontaneously mentioned and therefore not included in their group of canonical categories. More generally, these who, what, where, when, and how-type questions are the foundation of good story-telling and so it is not surprising that they constrain the structure of autobiographical event recall (Neisser, 1982). Second, longer delays between the event and the initial memory report often result in longer memory narratives (Lee & Brown, 2003). In part, this may be due to narrative conventions regarding shared information as Bohannon (1988) found that *time* information was much more likely to be included in reports obtained after three years than in those recorded after only one week.The frequency and timing of retrievals can also influence total narrative length (Nadel *et al*., 2007). Third, Marsh and Tversky (2004) have shown that the functional context of retrieval can influence what and how much individuals report. Similarly, reporting medium (e.g., verbal vs. typed) influences both absolute word count and the units of information provided in memory reports (Grysman & Denney, 2016). Therefore, elaboration is a less useful measure than one might expect.

Instead, defining vividness as a phenomenological experience of remembering perceptual detail can be informative. Rubin and Kozin (1984) tried to reframe FBMs as "vivid memories", as they thought that enhanced vividness was the defining feature of the phenomenon. In fact, FBMs often exhibit ceiling effects in vividness ratings regardless of the delay between event and memory report (Kvavilashvili *et al*., 2010; Kvavilashvili *et al*., 2003; Niedzwienska, 2003; Talarico & Rubin, 2003; Tinti *et al*., 2009; Weaver & Krug, 2004).Therefore, FBMs are more vivid than some ordinary memories (see Kvavilashvili *et al*., 2010 for direct comparison), but other types of AMs, such as those with high emotional intensity, are as vivid as FBMs (Talarico, LaBar, & Rubin, 2004). That is, the distributions of vividness overlap so there is no discontinuity that would require a special mechanism.

Confidence

In contrast to objective evidence of memory inaccuracy, participants consistently report enhanced confidence in FBM accuracy. FBMs are usually recalled with a higher degree of confidence than other memories of equal age (Brown & Kulik, 1977; Kvavilashvili *et al.*, 2010; Paradis, Solomon, Florer, & Thompson, 2004; Talarico & Rubin, 2003, 2006; Weaver, 1993; see Denver, Lane, and Cherry, 2010 for equally high confidence ratings in FBMs and ordinary AMs), even when individuals are confronted with evidence that the event in memory could not have occurred as it is remembered (Neisser & Harsch, 1992). Confidence is often at ceiling for FBMs (Christianson & Engelberg, 1999; Curci *et al.*, 2014; Denver, Lane, and Cherry, 2010; Neisser *et al.*, 1996; Niedzwienska, 2003; Talarico & Rubin, 2003, 2006; Weaver, 1993; Weaver & Krug, 2004) and often remains that high for at least months after the event (Christianson & Engelberg, 1999; Coluccia, Bianco, & Brandimonte, 2010; Conway *et al.*, 2009; Denver, Lane, and Cherry, 2010; Hirst *et al.*, 2015; Kraha & Boals, 2014; Kvavilashvili *et al.*, 2009; Liu, Ying, & Luo, 2012; Niedzwienska, 2003; Weaver & Krug, 2004). It may be that confidence ratings are based on equally reliably enhanced vividness ratings, as the two are correlated (Neisser & Harsch, 1992).

Therefore, along with vividness, the second distinctive property of FBMs is a discrepancy between metacognitive perception and objective reality. In fact, this discrepancy may have led to the identification of the phenomenon in the first place and may well lead to the most interesting applications of the phenomenon to ordinary memory processing. Thus, it is the secondary, phenomenological characteristics like vividness and confidence that may serve to retain the utility of FBM as a distinct category. If FBMs are differentiated by phenomenological experience, then the mechanisms responsible for the phenomenon must account for these differences, not explain (non-existent) encoding or retrieval differences.

The primary question in FBM research then becomes why we are more likely to maintain vivid, confidently held memories of these particular events. Enhancements to the subjective experience of remembering support Berntsen's (2009) model of FBM formation being driven by social group identification and a subsequent feedback loop where the memory itself serves to perpetuate identification with the social group.This may be considered a special case of Fernández's (2015) epistemic function of AMs more generally (e.g., an individual benefits from believing herself to be a patriot by relying on her vivid memory for learning about the September 11th attacks as evidence for that belief). However, we are still left to identify which event features are necessary to produce the defining memory characteristics.

Event conditions necessary to produce FBM

The vast majority of research in this field has been done in the aftermath of a public tragedy. This has been because consequentiality, distinctiveness, and emotional affect have been the primary features of the event thought to influence the formation of FBM (i.e., a memory report which satisfies the criteria described above for vividness and confidence, longevity and/or consistency). Here, we will discuss objective characteristics of the event thought to produce FBMs. Subjective assessments of the events will be discussed later.

Consequentiality

Consequential events most often studied include disasters with loss of life (e.g., earthquakes, terrorist attacks) or events with political implications (e.g., assassinations, resignations, invasions). FBM research is differentiated from traumatic memory research as the participants in the latter are directly affected by the events being studied. In the case of FBMs, participants are rarely so personally involved (see Pillemer, 2009 for a review of this distinction). However, the events being investigated are often on such a scale that the aftermath affects the lives of participants in other, more subtle ways. Is this comprehensive consequentiality responsible for FBMs? In short, no. The best evidence for the importance of consequentiality is indirect – an association between physical proximity to the event location and FBM formation. Cross-national studies have found this effect (Conway *et al.*, 1994; Curci *et al.*, 2001; Curci & Luminet, 2006; Kvavilashvili *et al.*, 2003) as have national studies with participants sampled from multiple locations (Er, 2003; Neisser *et al.*, 1996; Sharot *et al.*, 2007), although Luminet *et al.* (2004) found few differences in FBM specificity by nationality, and Pezdek (2003) found that those closest to the event were less likely to recall their personal circumstances than were those living far away. For consistency, some studies have found differences as a result of proximity (Conway *et al.*, 1994; Er, 2003) whereas others have not (Curci *et al.*, 2001; Curci & Luminet, 2006). Enhanced vividness, however, does seem to be associated with being physically closer to the event (Kvavilashvili *et al.*, 2003; Sharot *et al.*, 2007).

Many of the studies examining distance conflate directly experiencing the event with closest physical proximity. In cases involving political events (Conway *et al*., 1994; Curci *et al*., 2001), simply living in the affected nation ought to be less consequential than experiencing a natural disaster (Er, 2003; Neisser *et al*., 1996) or terrorist attack (Sharot *et al*., 2007), for example. The bulk of the empirical evidence fails to support the claim that objective consequentiality is relevant for the formation (Er, 2003; Tekcan, 2001), accuracy (Berntsen & Thomsen, 2005), consistency (Niedzwienska, 2003; Weaver, 1993), or vividness (Berntsen & Thomsen, 2005; Rubin & Kozin, 1984) of FBMs. However, Koppel, Brown, Stone, Coman, and Hirst (2013) showed that different factors predicted consistency for consequential (i.e., the first inauguration of U.S. President Obama) and non-consequential (i.e., the emergency landing of US Airways Flight 1549) public events; emotional intensity and significance predicted consistency of memories for a consequential event, whereas rehearsal predicted consistency for a non-consequential event. So, although (inter)national events may retain value by generating large numbers of potential participants, exclusively examining such events simply because they are assumed to have a requisite degree of consequentiality seems unnecessary.

Distinctiveness

The evidence in support of distinctiveness effects is much stronger than was found for consequentiality, as it has been correlated with the formation (Edery-Halpern & Nachson, 2004; Larsen, 1992; Wright & Gaskell, 1992) and vividness (Bohn & Berntsen, 2007; Edery-Halpern & Nachson, 2004) of FBMs. Mahmood, Manier, and Hirst (2004) found no relationship between distinctiveness and the formation or vividness of FBMs, but, distinctiveness in their study was defined as the first event in a series of similar, emotional, personally significant events (i.e., the deaths of lovers, friends, and/or family members due to AIDS). An event may be distinctive for reasons other than that it is the first of its kind, however; Brown and Kulik (1977) studied memory for a series of assassinations of political figures in a relatively brief period of time, yet each was a distinctive event. Edery-Halpern and Nachson (2004) found that the least distinctive event in their sample of terrorist attacks within a two-year period in Israel was also significantly less well remembered (i.e., fewer details were recalled and more responses were left blank for this memory compared to the other events). This is not surprising given that the episodic memory literature includes ample evidence of a distinctiveness advantage (i.e., *von Restorff effect*; see Schmidt, 2012 for a review within an FBM context). Brewer's (1988) study of AM suggests that the lower the frequency of event occurrence, the greater the likelihood of later cued recall. Thus, any effects of distinctiveness that might be present may be similar for FBM and AM.

Emotional Affect

Another event feature known to enhance ordinary memory and thought to influence FBMs is negative emotional affect. For example, negative stimuli "pop-out" in a neutral context to a greater extent than neutral stimuli in a fearful context (Ohman, Flykt, & Esteves, 2001). As they typically involve disasters, attacks, and assassinations, most FBM studies have included only negative events. Kraha, Talarico, and Boals (2014) examined the surprising announcement of Osama bin Laden's death, which was interpreted positively by the U.S. students sampled, and found little evidence of FBM. However, this lack of phenomenological enhancement was attributed more to lack of emotional intensity (see below) than to positive affect per se.

Other studies have shown both positive and negative events can lead to FBMs (Berntsen & Thomsen, 2005; Bohn & Berntsen, 2007; Demiray & Freund, 2015; Liu,Ying & Luo, 2012; Scott & Ponsoda, 1996;Tekcan, 2001). Positive interpretations of an event are more likely to produce FBMs (Bohn & Bernsten, 2007) and increase accuracy, vividness, and rehearsal of that event relative to the memories of those who interpret the event negatively (Breslin & Safer, 2011; Talarico & Moore, 2012). Kensinger and Schacter (2006) and Holland and Kensinger (2012) both found that consistency was higher (although confidence was lower) for positive interpretations of a sporting event and an election outcome, respectively. (However, in their examination of a sporting event, Talarico and Moore [2012] found no differences in consistency or confidence among fans of the winning and losing teams.) Generally, these results are consistent with pleasantness biases in autobiographical recall (see Walker, Skowronski, & Thompson, 2003 for a review). Furthermore, there is evidence from collective memories that even profoundly negative events are more likely to persist in the culture if they evoke positive connotations. For example, Hirst and Meksin (2009) describe how the assassinations of Lincoln and Kennedy endure because each President was subsequently deified by popular culture. Similarly, there is often an emphasis on patriotism and heroism in the face of tragedy (e.g., the Pearl Harbor or September 11th attacks) in societal recollections of those events (see Stone & Jay, Chapter 8, this volume, for further discussion of FBM for positive events). Summarizing the conditions necessary to produce FBMs, what could be a unique characteristic of FBMs (consequentiality) fails to predict the memory phenomenon, and well-characterized features of AM moderately account for the key characteristics of vividness and confidence, suggesting that subjective interpretations of events may be more responsible for the phenomena.

Individual processing of the subjective flashbulb event

Characteristics of how an individual processes the event at encoding and during rehearsal/retrieval are thought to be important determinants of FBMs. Encoding factors are closely tied to event features: distinctiveness with surprise, emotional affect with emotional intensity, and consequentiality with significance. What differentiates them is the objective vs. subjective nature of assessment. How these characteristics contribute to the FBM phenomenon individually and interactively is the focus of much current work in this area.

Significance

Significance, or personal importance, refers to the individual's subjective assessment of the event. This assessment can be based on any number of felt criteria and is not necessarily related to any material changes in the individuals' circumstances. As with consequentiality and emotion, group membership is frequently used as a proxy for significance. This provides a methodological advantage in identifying large numbers of participants for whom individual reactions to a common event are expected to vary. There are also theoretical reasons to justify this technique.

Berntsen (2009) argues quite persuasively that it is an event's importance to our social identity specifically that determines whether an event will produce an FBM. Because FBM research has emphasized recall of public events, it is not surprising that social identity is the most salient criteria for determining significance. Various traits have been used to demarcate social groups, including race (Brown & Kulik, 1977) gender, (Wright, Gaskell, & O'Muircheartaigh, 1998), language (Stone *et al*., 2013), religion (Curci *et al*., 2014; Lanciano, Curci, & Soleti, 2013; Tinti *et al*., 2009), and explicit membership in social movements (e.g., participation in resistance activity during military occupation by foreign forces (Berntsen & Thomsen, 2005). Across all operational definitions, social group membership (and therefore personal significance) was positively associated with FBM formation (Brown & Kulik, 1977;Wright *et al*., 1998), enhanced accuracy (Berntsen & Thomsen, 2005), greater consistency (Curci *et al*., 2014; Lanciano *et al*., 2013; Tinti *et al*., 2009), increased elaboration (Curci *et al*., 2014; Stone *et al*., 2013), and enhanced vividness (Berntsen & Thomsen. 2005; Tinti *et al*., 2009), though it was unrelated to confidence (Curci *et al*., 2014).

Lanciano *et al*. (2013) further showed that ratings of importance (a conflation of event consequentiality and personal significance) for the death of Pope John Paul II differed as a function of religious affiliation, consistent with Berntsen's (2009) social identity model. Similarly, Tinti *et al*. (2009) used national origin as an alternate avenue to identification with the Pontiff, allowing it to serve as a proxy for significance. After controlling for religiosity, they found that participants with no particular affiliation to Pope John Paul II (i.e., Swiss individuals) had less consistent and less vivid FBMs for his death than did participants who could identify with him (i.e., Italian and Polish individuals). Luminet *et al*. (2004) showed that U.S. versus non-U.S. citizenship status was more predictive of developing an FBM for the September 11th attacks than was nation of residence. In Stone *et al*.'s (2013) examination of linguistic groups within Belgium, both French and Dutch speakers thought the division of the University of Leuven was equally important, but Dutch speakers were more likely to be personally and/or politically involved in the division and, as a result, had more elaborate FBMs and rehearsed those memories more frequently than did French speakers.

Not all FBM studies rely on group membership to examine this characteristic; individual ratings of personal importance have also been investigated (when sufficient variability exists to do so). Participant ratings of significance are positively correlated with FBM formation (Bohannon & Symons, 1992; Conway *et al.*, 1994; Larsen, 1992; Mahmood, Manier, & Hirst, 2004; Wright & Gaskell, 1992; but see Wright *et al.*, 1998). Paradis *et al.* (2004) found that their New York City participants rated both September 11th and 12th as personally important and their sample developed FBMs for both of those days, in terms of initial recall and later consistency. Niedzwienska (2003) also found significance to be correlated with consistency of the FBM report. Furthermore, personal significance has been reliably related to vividness ratings (Mahmood, Manier, & Hirst, 2004; Nachson & Zelig, 2003; Niedzwienska, 2003; Rubin & Kozin, 1984) and to the number of details reported (i.e., elaboration) in German, Turkish, British, and American samples, though not within a Chinese sample (Kulkofsky *et al.*, 2011). The authors attributed this anomaly to collectivist cultural expectations to deemphasize personal goals and activities that therefore potentially dampened significance ratings within this group.

However, individual ratings of personal significance are not universally found to predict FBM phenomena. Otani *et al.* (2005) classified memory reports as FBMs or non-FBMs, yet found no difference in the significance ratings of participants in each group. Davidson and Glisky (2002) also found no differences in the signifi- cance ratings of two events, yet one event led to reliably more FBM reports than the other.

As the study of irrelevant characteristics (e.g., consequentiality, negative valence) declines and is replaced by more systematic study of relevant characteristics, the nuanced nature of such effects can be determined. It is our belief that personal significance as assessed by social group membership will remain a determining feature of FBM formation, but that other criteria for personal significance may also lead to vivid, confidently held FBMs as well. As with distinctiveness, there is an abundance of data for a self-referential effect in memory performance, with personally relevant material enhancing memory (see Symons & Johnson, 1997 for a review). Thus, the influence of significance on FBMs can be predicted from more general features of AMs.

Surprise

As significance is differentiated from consequentiality, so, too, is surprise different from distinctiveness. Surprise is a personal, emotional reaction to the event, not a property of the event. Note that although an event can be expected, and therefore not surprising, it can still be distinctive, as was seen in the case of several terrorist attacks in Israel, the sad inevitability of which does not prevent each attack from being distinct (Edery-Halpern & Nachson, 2004).

Although surprise is a key component in many models of FBM formation (Brown & Kulik, 1977; Er, 2003; Finkenauer *et al.*, 1998), there is little data to support its inclusion. Only when comparing FBM for two similar events (the deaths of Princess Diana and Mother Theresa), were higher ratings of surprise associated with greater initial recall and later consistency of FBM. Surprise only seems to be directly influential when tied to social identity, otherwise it plays no direct role in FBM formation (Curci & Luminet, 2009).

FBMs have been found for expected events (Curci *et al.*, 2001; Lanciano *et al.*, 2013; Tekcan, 2001; Tinti *et al.*, 2009; Winograd & Killinger, 1983). Additionally, Coluccia, Bianco, & Brandimonte (2010) found no differences between expected and unexpected events in the relationships between delay and consistency, confidence, or elaboration. Equal surprise ratings were provided by those who did and those who did not develop FBMs for the Kobe earthquake (Otani *et al.*, 2005). There is even evidence that surprise and FBM phenomena are negatively correlated. Berntsen and Thomsen (2005) found that participants rated the invasion of Denmark as more surprising than its liberation, but were more likely to have FBMs for the liberation than for the invasion. It may be that surprise is a retrospective evaluation that is based on the emotional intensity of one's reaction. If this is the case, we would expect surprise to play a role in FBM formation when it is present, but it would be neither a necessary nor a sufficient condition for the phenomenon. Interestingly, the effect of surprise in AM has not been thoroughly investigated.

We do know that surprise tends to be a positive emotion in ordinary AMs (Talarico, LaBar, & Rubin, 2004). In other words, when cued to generate memories of sur- prise, participants are more likely to recall pleasant events (e.g., a surprise birthday party) than unpleasant events (e.g., an unexpectedly low score on an exam). Yet, when evaluating positive and negative events generally, positive events tend to be expected (e.g., weddings) whereas negative events are typically unexpected (e.g., divorce) (Berntsen, 2002; Rubin & Berntsen, 2003). The public events studied in FBM research are predominantly surprising and negative, therefore the scope of FBM research has been limited by what may be coincidental, not causal, event features. Anticipatable public events (e.g., elections) may be ripe for additional exploitation by FBM investigators (e.g., Boals, 2010; Holland & Kensinger, 2012).

Emotional Intensity

Some of the most contradictory findings in the FBM literature are those involving emotional intensity. The positive-negative valence dimension depends primarily on the nature of the event itself (though some events may be interpreted differently among different groups). Depth of feeling, however, is necessarily a subjective reaction to a given event within a valence category.

There are data supporting emotional intensity's role in FBM formation using participant ratings (Berntsen & Thomsen, 2005; Bohannon, 1988; Bohannon & Symons, 1992; Davidson & Glisky, 2002; Paradis *et al.*, 2004) and using culture as a proxy for emotion (Brown & Kulik, 1977; Curci *et al.*, 2001), but nearly as many studies that fail to find a correlation with participant ratings (Otani *et al.*, 2005; Smith, Bibi, & Sheard, 2003; Tekcan, 2001), or with culture as a proxy (Luminet *et al.*, 2004).Wright *et al.* (1998) found that men, who rated the Hillsborough football disaster as less emotional than women, were more likely to develop FBM, not less. The same contradictory pattern emerges for intensity and consistency, with some studies finding a positive relationship between the two (Bohannon & Symons, 1992; Conway *et al.*, 1994; Davidson & Glisky, 2002; Schmolck, Buffalo, & Squire, 2000), but more that fail to find such a relationship (Christianson & Engelberg, 1999; Kvavilashvili *et al.*, 2009; Nachson & Zelig, 2003; Neisser *et al.*, 1996; Neisser & Harsch, 1992; Schmidt, 2004; Talarico & Rubin, 2003). Vividness of the FBM is equally divergent. Rubin and Kozin (1984) failed to find a correlation between emotional intensity and vividness, whereas others have found the two to be related (Berntsen & Thomsen, 2005; Nachson & Zelig, 2003; Pillemer, 1984). Lastly, Hirst *et al.* (2015) failed to find a relationship between emotional intensity and confidence at either three-year or ten-year intervals. Even if the effects of emotional intensity were more reliable, the advantage of emotionally intense experiences over neutral events is well established in the episodic and AM literatures and so it cannot be seen as unique to FBM.

Rehearsal

The final processing feature we will examine is the only non-encoding-specific mechanism discussed in this literature. The effects of rehearsal on FBM seem to dissociate based on the dependent variable of interest. Increased rehearsal has been correlated with the formation of FBMs (Bohannon, 1988; Bohannon & Symons, 1992; Curci *et al.*, 2001; Davidson & Glisky, 2002; Otani *et al.*, 2005; Tekcan & Peynircioglu, 2002; Tinti *et al.*, 2014), although Hornstein, Brown, and Mulligan (2003) found that to be true only for covert, not overt, rehearsal. For vividness, there seems to be no relationship with rehearsal (Pillemer, 1984; Rubin & Kozin, 1984). Hirst *et al.* (2015) found that confidence was correlated with rehearsal after a three-year delay (as did Kvavilashvili *et al.*, 2009), but not after a ten-year delay.

For consistency, the pattern is quite variable. There are studies showing a positive correlation between rehearsal and consistency (Bohannon & Symons, 1992; Conway *et al.*, 2009; Davidson & Glisky, 2002; Schmolck, Buffalo, & Squire, 2000), a negative correlation between the two (Kvavilashvili *et al.*, 2009), and no relationship (Kvavilashvili *et al.*, 2009; Pillemer, 1984; Schmolck, Buffalo, & Squire, 2000), across delays ranging from six months to three years. There is no systematic relationship among delay, type of rehearsal (overt vs. covert), or consistency.

Moreover, rehearsal via media exposure can increase errors of commission as is seen in the "crashing memories" phenomenon (Crombag, Wagenaar, & van Koppen, 1996) of reporting false memories of non-existent footage (Ost *et al.*, 2008). Media exposure can also serve to increase omissions via retrieval-induced forgetting (Coman, Manier, & Hirst, 2009). These details are unlikely to be reintroduced into the memory narrative because there are few opportunities for correction. As with most AMs, there is little evidence of what actually occurred, therefore little can be presented to refute or corroborate one's personal recollection.

Yet, we know that rehearsal is a potent mechanism for sustaining memory. Repeated overt retrieval attempts lead to more consistent and more elaborated AM reports (Campbell *et al.*, 2011; Nadel *et al.*, 2007). In other words, individuals add details to their memory reports while also maintaining the originally provided information. Kvavilashvili *et al.* (2010) found a positive correlation between rehearsal and consistency for a staged autobiographical event in contrast to a nega- tive correlation for FBM of the September 11th attacks. Specifying the differences between public and private, overt and covert rehearsal behaviors may help make sense of these disparate results.

Combining factors

Most events are chosen as subjects of FBM research because they exhibit many of the features we have noted here. This has led some investigators to adopt statistical techniques such as latent variable modeling and structural equation modeling to determine the relationships among these features (see Luminet, Chapter 3, this volume) frequently within FBM and non-FBM of the same event. Each of these models defined FBM as recall of some number of canonical categories or some measure of completeness and specificity, not vividness or confidence, the more reliable characteristics to differentiate FBMs from ordinary AMs. Further, these models consider significance and distinctiveness to be indirect factors in FBM formation, though the data presented here suggest that they are among the predominant mechanisms responsible for FBM.

Day and Ross (2014) are the only investigators to model confidence, one of the more reliable FBM characteristics. They found that the strongest predictor of confidence when remembering Michael Jackson's death was a sense of attachment to the performer (which lead to greater initial surprise, increased emotional intensity, and enhanced rehearsal). Furthermore, they found that individuals' beliefs about the persistence of FBM also correlated with their later confidence in their own memories, but not with the consistency of those memory reports. Models like this suggest a high degree of interrelatedness among metacognitive features (i.e., beliefs about memory generally and beliefs about one's own memory) and between the ways in which events are processed and their resulting phenomenology. In contrast, there seems to be little predictive value of event features on resulting mnemonic characteristics.

Summary and future directions

FBMs are distinguished from ordinary memories by their vividness and the confidence with which they are held. There is little evidence that they are reliably different from ordinary AMs in longevity, accuracy, or consistency. Features of the event and processing characteristics of the individual identifiable in ordinary AM explain enhancements to each of these memory properties. Curci and Lanciano (2009), Lanciano and Curci, (2012; see also chapter 1, this volume) have argued that FBMs are better viewed as a category separate from ordinary memories rather than along a continuum with them. To support this claim they conducted two studies, each with a large number of participants who were asked to recall details of a single public event on two occasions. The consistency of their responses to canonical FBM questions were rated and analyzed statistically. In both studies, the authors found a better fit of the empirical data to a categorical (taxonomic) model than to a continuous (dimensional) model of memory.

Further, guidance for interpreting the claim of distinct categories is based on a limited theory of episodic/autobiographical memory. Specifically, Conway (1995) has argued that FBMs are stable clusters of sensory-perceptual details that are highly integrated and therefore accessed holistically. They are contrasted with ordinary AMs which are dynamic reconstructions of information drawn from event-specific knowledge and generalized autobiographical themes. Brewin (2014) reinforces this idea when suggesting that FBMs are defined by "a more detailed perceptual record of experience than ordinary memories" (p. 73). He goes on to argue that FBMs are evidence of a distinct perceptual memory system, separate from episodic memory, and particularly linked with emotional reactivity. It is this specific, empirically unsupported difference between FBM and AM that is discussed in both papers, and thus is not an ideal assumption for testing the nature of categories.

A more common and better supported view of AM organization is based on a more flexible view of the construction of autobiographical memories at all levels (e.g., Bartlett, 1932; Hirst & Phelps, 2016; Neisser, 1982; Rubin, 2012; Rubin & Umanath, 2015). This view allows for more flexible use of non-hierarchical knowledge bases and can adequately account for the findings reviewed in our chapter for both FBM and ordinary AM. For the event conditions, consequentiality seems to be irrelevant to FBM, distinctiveness is the most predictive, and emotional affect is as yet understudied. Of the processes discussed, significance is the most promising determinant of FBM, especially as related to social group membership in the context of public events. It is correlated with formation, consistency, vividness, and confidence of FBMs. Surprise has an unreliable influence on the formation and consistency of FBMs (the least reliable features of FBMs overall) and its effects on vividness and confidence (the most reliable features) have not been systematically investigated.Therefore, this seems to be one of the more promising areas for future investigation. Emotional intensity is an unreliable predictor of FBM, yet this is most likely due to inconsistencies in defining FBM. At least some of the enhanced confidence and/or vividness associated with FBMs are probably due to enhanced emotionality; however, the exact nature and scope of that influence has yet to be determined. Lastly, rehearsal tends to be correlated with the formation of FBMs, but the relationships between rehearsal and consistency, vividness, and confidence are quite variable. Because rehearsals can take multiple forms (public vs. private, overt vs. covert) and can simultaneously enhance recall of some details while introducing erroneous details, disentangling the specific effects of rehearsal on particular features of FBM is an important goal for future investigators.

We have no doubt that FBM will remain a topic of frequent investigation. After 40 years of wandering, we may be finally approaching the promised land of under- standing this fascinating and complex phenomenon. Berntsen's (2009) resurrection and expansion of the social aspects of Brown and Kulik's (1977) hypothesis may be the most significant contribution to the FBM literature in recent years, in part because it did not rely on any claims to special mechanisms; these factors are important to all AMs. By reclaiming the public nature of events from mere methodological convenience to an important theoretical feature, Berntsen's (2009) model helps to sustain FBM as a distinctive phenomenon and as a fruitful area for inter- and multi-disciplinary work on remembering.

Listening to news is not an inherently memorable activity (Larsen, 1992). The counterintuitive enhancement of memory for otherwise mundane activities was what spurred initial interest in FBM. Neisser's (1982) suggestion that "we remember the details of a flashbulb occasion because those details are the links between our own histories and History" (p. 48) was prescient, but perhaps too broad. Evidence of FBM phenomena from events that are public, but on a relatively smaller scale than is typical of natural disasters or political events, are proof of this. For example, memories of sporting events for fans of a given team underscore the role of personal significance (and the irrelevance of objective consequentiality) in Flashbulb memory. These memories are likely only among those who consider the event to be important, and therefore emotionally intense. Groups of fans observing subsequent games serve as a natural context for rehearsal and sharing one's memory serves as a potent symbol of one's identity as a "true" fan. Rehearsing one's personal relationship to an event rather than the factual details of the event itself is consistent with the role of the feedback loop in creating and maintaining FBMs. Telling one's personal story of learning about a shared event serves narrative conventions to introduce novel details to a conversation and serves larger social functions by building group cohesion. The expectation of remembering, of having that story to share, is an underappreciated characteristic of the FBM phenomenon. Both individuals and groups have that expectation. There is presumed judgment of one's failure to remember a given event and what that failure implies about the individual's understanding or interpretation of a given event.

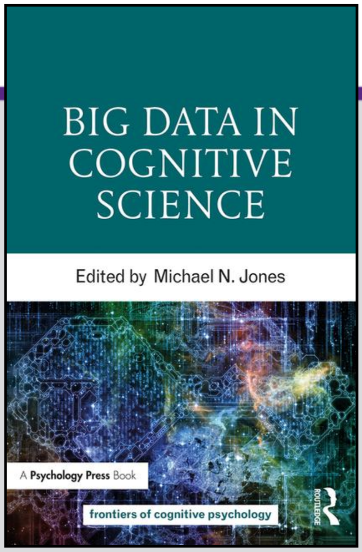
One important direction for future research ought to be moving away from indirect assessment of social identity (e.g., cross-cultural studies) and toward a priori measures of group identification and selection of samples who vary significantly on those measures. One should predict that FBM phenomenology is likely to be greatest when membership in a social group is important (not just incidental) to an individual and when that membership is vulnerable. Because every individual is simultaneously a member of multiple social groups, each of those groups is not of equal importance. Feelings of insecurity about one's group membership should enhance the perceived significance of both the event itself and of the FBM for its symbolic function. The pressure to prove one's status should correlate with phenomenological features of the memory such as vividness and confidence in its accuracy. For groups wherein one's status is more established or more assured, the FBM phenomenology may still be at the higher end of a continuum within AM, but they may not be at the extreme ceiling seen for other events.

In his book, *Extraordinary memories for exceptional events*, Schmidt (2012) defines FBM methodologically (i.e., studies asking individuals "how did you first learn the news of public event X"), phenomenologically (i.e., as a detailed, long-lasting, vivid, and confidently-held AM), and theoretically (i.e., accurate and detailed memories that are the result of a special remembering mechanism). We argue here that the theoretical definition is unsustainable. There is insufficient evidence to support a categorical dissociation between FBMs and other AMs. The phenomenological definition is nearing refinement but there are still some unanswered questions. We believe this to be the most fruitful and therefore most important direction for FBM research in the coming years. Most importantly, we believe that a clear, *a priori* definition of what constitutes an FBM or, better, a set of alternative definitions that can be contrasted to see which provides a better understanding of the data, is necessary in future empirical work. Clearly stating what data are required to identify a given memory as an FBM or a non-FBM will increase the utility of new studies in addressing these important questions. Without such definitions, future studies risk simply repeating past studies on new events and merely describing the phenomenon rather than explaining it. We have also argued that the methodological definition may be more than a simple convenience. The personal memory of learning about a public event remains a unique domain within AM. This characteristic ties event features to memory features and therefore provides clues to potential mechanism(s) for the phenomenon. It also captures why the phenomenon is of such enduring interest. Long-lasting, detailed, vivid, confidently-held memories of directly-experienced, emotional, significant, and well-rehearsed events is not unexpected, but the idea that simply receiving news can change a mundane experience into a noteworthy memory remains surprising and is why FBMs deserve to remain a distinct phenomenon.

CHAPTER

5

DEVELOPING COGNITIVE THEORY BY MINING LARGE-SCALE NATURALISTIC DATA



This chapter is excerpted from

Big Data in Cognitive Science

edited by Michael N. Jones.

Learn more

DEVELOPING COGNITIVE THEORY BY MINING LARGE-SCALE NATURALISTIC DATA

Excerpted from *Big Data in Cognitive Science*

Abstract

Cognitive research is increasingly coming out of the laboratory. It is becoming much more common to see research that repurposes large-scale and naturalistic data sources to develop and evaluate cognitive theories at a scale not previously possible. We now have unprecedented availability of massive digital data sources that are the product of human behavior and offer clues to understand basic principles of cognition. A key challenge for the field is to properly interrogate these data in a theory-driven way to reverse engineer the cognitive forces that generated them; this necessitates advances in both our theoretical models and our methodological techniques. The arrival of Big Data has been met with healthy skepticism by the field, but has also been seen as a genuine opportunity to advance our understanding of cognition. In addition, theoretical advancements from Big Data are heavily intertwined with new methodological developments—new techniques to answer questions from Big Data also give us new questions that could not previously have been asked. The goal of this volume is to present emerging examples from across the field that use large and naturalistic data to advance theories of cognition that would not be possible in the traditional laboratory setting.

While laboratory research is still the backbone of tracking causation among behavioral variables, more and more cognitive research is now letting experimental control go in favor of mining large-scale and real-world datasets. We are seeing an exponential expansion of data available to us that is the product of human behavior: Social media, mobile device sensors, images, RFID tags, linguistic corpora, web search logs, and consumer product reviews, just to name a few streams. Since 2012, about 2.5 exabytes of digital data are created every day (McAfee, Brynjolfsson,

DEVELOPING COGNITIVE THEORY BY MINING LARGE-SCALE NATURALISTIC DATA

Excerpted from *Big Data in Cognitive Science*

Davenport, Patil, & Barton, 2012). Each little piece of data is a trace of human behavior and offers us a potential clue to understand basic cognitive principles; but we have to be able to put all those pieces together in a reasonable way. This approach necessitates both advances in our theoretical models and development of new methodological techniques adapted from the information sciences.

Big Data sources are now allowing cognitive scientists to evaluate theoretical models and make new discoveries at a resolution not previously possible. For example, we can now use online services like Netflix, Amazon, and Yelp to evaluate theories of decision-making in the real world and at an unprecedented scale. Wikipedia edit histories can be analyzed to explore information transmission and problem solving across groups. Linguistic corpora allow us to quantitatively evaluate theories of language adaptation over time and generations (Lupyan & Dale, 2010) and models of linguistic entrainment (Fusaroli, Perlman, Mislove, Paxton, Matlock, & Dale, 2015). Massive image repositories are being used to advance models of vision and perception based on natural scene statistics (Griffiths, Abbott, & Hsu, 2016; Khosla, Raju, Torralba, & Oliva, 2015). Twitter and Google search trends can be used to track the outbreak and spread of "infectious" ideas, memory contagion, and information transmission (Chen & Sakamoto, 2013; Masicampo & Ambady, 2014; Wu, Hofman, Mason, & Watts, 2011). Facebook feeds can be manipulated to explore information diffusion in social networks (Bakshy, Rosenn, Marlow, & Adamic, 2012; Kramer, Guillory, & Hancock, 2014). Theories of learning can be tested at large scales and in real classroom settings (Carvalho, Braithwaite, de Leeuw, Motz, & Goldstone, 2016; Fox, Hearst, & Chi, 2014). Speech logs afford both theoretical advancements in auditory speech processing, and practical advancements in automatic speech comprehension systems.

DEVELOPING COGNITIVE THEORY BY MINING LARGE-SCALE NATURALISTIC DATA

Excerpted from *Big Data in Cognitive Science*

The primary goal of this volume is to present cutting-edge examples that use large and naturalistic data to uncover fundamental principles of cognition and evaluate theories that would not be possible without such scale. A more general aim of the volume is to take a very careful and critical look at the role of Big Data in our field. Hence contributions to this volume were handpicked to be examples of advancing theory development with large and naturalistic data.

What is Big Data?

Before trying to evaluate whether Big Data could be used to benefit cognitive science, a very fair question is simply *what is Big Data*? Big Data is a very popular buzzword in the contemporary media, producing much hype and many misconceptions. Whatever Big Data is, it is having a revolutionary impact on a wide range of sciences, is a "game-changer," transforming the way we ask and answer questions, and is a must-have for any modern scientist's toolbox. But when pressed for a definition, there seems to be no solid consensus, particularly among cognitive scientists. We know it probably doesn't fit in a spreadsheet, but opinions diverge beyond that. The issue is now almost humorous, with Dan Ariely's popular quip comparing Big Data to teenage sex, in that "everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it."

As scientists, we are quite fond of careful operational definitions. However, Big Data and data science are still-evolving concepts, and are moving targets for formal definition. Definitions tend to vary depending on the field of study. A strict interpretation of Big Data from the computational sciences typically refers to datasets that are so massive and rapidly changing that our current data processing methods are inadequate. Hence, it is a drive for the development of distributed storage platforms and algorithms to analyze datasets that are currently out of reach. The term extends to challenges inherent in data capture, storage, transfer, and predictive analytics. As a loose quantification, data under this interpretation currently become "big" at scales north of the exabyte.

Under this strict interpretation, work with true Big Data is by definition quite rare in the sciences; it is more development of architectures and algorithms to manage these rapidly approaching scale challenges that are still for the most part on the horizon (NIST Big Data Working Group, 2014). At this scale, it isn't clear that there are any problems in cognitive science that are true Big Data problems yet. Perhaps the largest data project in the cognitive and neural sciences is the Human Connectome Project (Van Essen et al., 2012), an ambitious project aiming to construct a network map of anatomical and functional connectivity in the human brain, linked with batteries of behavioral task performance. Currently, the project is approaching a petabyte of data. By comparison, the Large Hadron Collider project at CERN records and stores over 30 petabytes of data from experiments each year.

More commonly, the *Gartner 3 Vs* definition of Big Data is used across multiple fields: "Big data is high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision-making, insight discovery and process optimization" (Laney, 2012). *Volume* is often indicative of the fact that Big Data records and observes everything within a recording register, in contrast to our commonly used methods of sampling in the behavioral sciences. *Velocity* refers to the characteristic that Big Data is often a real-time stream of rapidly captured data. The final characteristic, *variety*, denotes that Big

DEVELOPING COGNITIVE THEORY BY MINING LARGE-SCALE NATURALISTIC DATA

Data draws from multiple qualitatively different information sources (text, audio, images, GPS, etc.), and uses joint inference or fusion to answer questions that are not possible by any source alone. But far from being expensive to collect, Big Data is usually a natural byproduct of digital interaction.

So while a strict interpretation of Big Data puts it currently out of reach, it is simultaneously everywhere by more liberal interpretations. Predictive analytics based on machine learning has been hugely successful in many applied settings (see Hu, Wen, & Chua, 2014, for a review). Newer definitions of Big Data summarize it as more focused on repurposing naturalistic digital footprints; the size of "big" is relative across different fields (NIST Big Data Working Group, 2014). The NIH BD2K (Big Data to Knowledge) program is explicit that a Big Data approach is best defined by what is large and naturalistic to specific subfields, not an absolute value in bytes. In addition, BD2K notes that a core Big Data problem involves joint inference across multiple databases. Such combinatorial problems are clearly Big Data, and are perfectly suited for theoretically driven cognitive models—many answers to current theoretical and practical questions may be hidden in the complimentary relationship between data sources.

What is Big Data to Cognitive Science?

Much of the publicity surrounding Big Data has focused on its insight power for business analytics. Within the cognitive sciences, we have been considerably more skeptical of Big Data's promise, largely because we place such a high value on explanation over prediction. A core goal of any cognitive scientist is to fully understand the system under investigation, rather than being satisfied with a simple descriptive or predictive theory.

DEVELOPING COGNITIVE THEORY BY MINING LARGE-SCALE NATURALISTIC DATA

Excerpted from *Big Data in Cognitive Science*

Understanding the mind is what makes an explanatory cognitive model distinct from a statistical predictive model—our parameters often reflect hypothesized cognitive processes or representations (e.g. attention, memory capacity, decision thresholds, etc.) as opposed to the abstract predictive parameters of, say, weights in a regression model. Predictive models are able to make predictions of new data provided they are of the same sort as the data on which the model was trained (e.g. predicting a new point on a forgetting curve). Cognitive models go a step further: An explanatory model should be able to make predictions of how the human will behave in situations and paradigms that are novel and different from the situations on which the model was built but that recruit the same putative mechanism(s) (e.g. explaining the process of forgetting).

Marcus and Davis (2014) have argued rather convincingly that Big Data is a scientific idea that should be retired. While it is clear that large datasets are useful in discovering correlations and predicting common patterns, more data do not on their own yield explanatory causal relationships. Big Data and machine learning techniques are excellent bedfellows to make predictions with greater fidelity and accuracy. But the match between Big Data and cognitive models is less clear; because most cognitive models strive to explain causal relationships, they may be much better paired with experimental data, which shares the same goal. Marcus and Davis note several ways in which paying attention to Big Data may actually lead the scientist astray, compared to a much smaller amount of data from a well-controlled laboratory scenario.

In addition, popular media headlines are chock-full of statements about how theory is obsolete now that Big Data has arrived. But theory is a simplified model of empirical phenomena—*theory explains data*. If anything, cognitive theory is more necessary to help us understand Big Data in a principled way given that much of the data were generated by the cognitive systems that we have carefully studied in the laboratory, and cognitive models help us to know where to search and what to search for as the data magnitude grows.

Despite initial skepticism, Big Data has also been embraced by cognitive science as a genuine opportunity to develop and refine cognitive theory (Griffiths, 2015). Criticism of research using Big Data in an atheoretical way is a fair critique of the way some scientists (and many outside academia) are currently using Big Data. However, there are also scientists making use of large datasets to test theory-driven questions—questions that would be unanswerable without access to large naturalistic datasets and new machine learning approaches. Cognitive scientists are, by training, [experimental] control freaks. But the methods used by the field to achieve laboratory control also serve to distract it from exploring cognitive mechanisms through data mining methods applied Big Data.

Certainly, Big Data is considerably more information than we typically collect in a laboratory experiment. But it is also naturalistic, and a footprint of cognitive mechanisms operating in the wild (see Goldstone & Lupyan, 2016, for a recent survey). There is a genuine concern in the cognitive sciences that many models we are developing may be overfit to specific laboratory phenomena that neither exist nor can be generalized beyond the walls of the lab. The standard cognitive experiment takes place in one hour in a well-controlled setting with variables that normally covary in the real world held constant. This allows us to determine conclusively that the flow of causation is from our manipulated variable(s) to the dependent variable, and often by testing discrete settings ("factorology"; Balota, Yap, Hutchison, & Cortese, 2012).

DEVELOPING COGNITIVE THEORY BY MINING LARGE-SCALE NATURALISTIC DATA

Excerpted from *Big Data in Cognitive Science*

It is essential to remember that the cognitive mechanisms we study in the laboratory evolved to handle real information-processing problems in the real world. By "capturing" and studying a mechanism in a controlled environment, we risk discovering experiment or paradigm-specific strategies that are a response to the experimental factors that the mechanism did not evolve to handle, and in a situation that does not exist in the real world. While deconfounding factors is an essential part of an experiment, the mechanism may well have evolved to thrive in a rich statistically redundant environment. In this sense, cognitive experiments in the lab may be somewhat analogous to studying captive animals in the zoo and then extrapolating to behavior in the wild.

The field has been warned about over-reliance on experiments several times in the past. Even four decades ago Estes (1975) raised a concern in mathematical psychology that we may be accidentally positing mechanisms that apply only to artificial situations, and that our experiments may unknowingly hold constant factors that may covary to produce very different behavior in the real world. More recently, Miller (1990) reminded cognitive scientists of Estes' reductionism caution:

> I have observed over the years that there is a tendency for even the best cognitive scientists to lose sight of large issues in their devotion to particular methodologies, their pursuit of the null hypothesis, and their rigorous efforts to reduce anything that seems interesting to something else that is not. An occasional reminder of why we flash those stimuli and measure those reaction times is sometimes useful.
>
> (Miller, 1990: 7)

DEVELOPING COGNITIVE THEORY BY MINING LARGE-SCALE NATURALISTIC DATA

Excerpted from *Big Data in Cognitive Science*

Furthermore, we are now discovering that much of the behavior we want to use to make inferences about cognitive mechanisms is heavy-tail distributed (exponential and power-law distributions are very common in cognitive research). Sampling behavior in a one-hour lab setting is simply insufficient to ever observe the rare events that allow us to discriminate among competing theoretical accounts. And building a model from the center of a behavioral distribution may fail horribly to generalize if the tail of the distribution is the important characteristic that the cognitive mechanism evolved to deal with.

So while skepticism about Big Data in cognitive science is both welcome and warranted, the above points are just a few reasons why Big Data could be a genuine opportunity to advance our understanding of human cognition. If dealt with in a careful and theoretically driven way, Big Data offers us a completely new set of eyes to understand cognitive phenomena, to constrain among theories that are currently deadlocked with laboratory data, to evaluate generalizability of our models, and to have an impact on the real-world situations that our models are meant to explain (e.g. by optimizing medical and consumer decisions, information discovery, education, etc.). And embracing Big Data brings with it development of new analytic tools that also allow us to ask new theoretical questions that we had not even considered previously.

How is Cognitive Research Changing with Big Data?

Cognitive scientists have readily integrated new technologies for naturalistic data capture into their research. The classic cognitive experiment typically involved a single subject in a testing booth making two alternative forced choice responses to stimuli presented on a monitor. To be clear, we have learned a great deal about fundamental principles of human cognition with this basic laboratory approach. But the modern cognitive experiment may involve mobile phone games with multiple individuals competing in resource sharing simultaneously from all over the world (Dufau et al., 2011; Miller, 2012), or dyads engaged in real-time debate while their attention and gestures are captured with Google Glass (Paxton, Rodriguez, & Dale, 2015).

In addition, modern cognitive research is much more open to mining datasets that were created for a different purpose to evaluate the models we have developed from the laboratory experiments. Although the causal links among variables are murkier, they are still possible to explore with new statistical techniques borrowed from informatics, and the scale of data allows the theorist to paint a more complete and realistic picture of cognitive mechanisms. Furthermore, online labor markets such as Amazon's Mechanical Turk have accelerated the pace of experiments by allowing us to conduct studies that might take years in the laboratory in a single day online (Crump, McDonnell, & Gureckis, 2013; Gureckis et al., 2015).

Examples of new data capture technologies advancing our theoretical innovations are emerging all over the cognitive sciences. Cognitive development is a prime example. While development unfolds over time, the field has traditionally been reliant on evaluating infants and toddlers in the laboratory for short studies at regular intervals across development. Careful experimental and stimulus control is essential, and young children can only provide us with a rather limited range of response variables (e.g., preferential looking and habituation paradigms are very common with infants).

DEVELOPING COGNITIVE THEORY BY MINING LARGE-SCALE NATURALISTIC DATA

While this approach has yielded very useful information about basic cognitive processes and how they change, we get only a small snapshot of development. In addition, the small scale is potentially problematic because many theoretical models behave in a qualitatively different way depending on the amount and complexity of data (Frank, Tenenbaum, & Gibson, 2013; McClelland, 2009; Qian & Aslin, 2014; Shiffrin, 2010). Aslin (2014) has also noted that stimulus control in developmental studies may actually be problematic. We may be underestimating what children can learn by using oversimplified experimental stimuli: These controlled stimuli deconfound potential sources of statistical information in learning, allowing causal conclusions to be drawn, but this may make the task much more difficult than it is in the real world where multiple correlated factors offer complimentary cues for children to learn the structure of the world (see Shukla, White, & Aslin, 2011). The result is that we may well endorse the wrong learning model because it explains the laboratory data well, but is more complex than is needed to explain learning in the statistically rich real world.

A considerable amount of developmental research has now come out of the laboratory. Infants are now wired with cameras to take regular snapshots of the visual information available to them across development in their real world experiences (Aslin, 2009; Fausey, Jayaraman, & Smith, 2016; Pereira, Smith, & Yu, 2014). LENATM recording devices are attached to children to record the richness of their linguistic environments and to evaluate the effect of linguistic environment on vocabulary growth (VanDam et al., 2016; Weisleder & Fernald, 2013). In one prominent early example, the SpeechHome project, an entire house was wired to record 200,000+ hours of audio and video from one child's first three years of life (Roy, Frank, DeCamp, Miller, & Roy, 2015).

DEVELOPING COGNITIVE THEORY BY MINING LARGE-SCALE NATURALISTIC DATA

Excerpted from *Big Data in Cognitive Science*

Tablet-based learning games are now being designed to collect theoretically constraining data as children are playing them all over the world (e.g. Frank, Sugarman, Horowitz, Lewis, & Yurovsky, 2016; Pelz, Yung, & Kidd, 2015).

A second prime example of both new data capture methods and data scale advancing theory is in visual attention. A core theoretical issue surrounds identification performance as a function of target rarity in visual search, but the number of trials required to get stable estimates in the laboratory is unrealistic. Mitroff et al. (2015) opted instead to take a Big Data approach to the problem by turning visual search into a mobile phone game called "Airport Scanner." In the game, participants act the part of a TSA baggage screener searching for prohibited items as simulated luggage passes through an x-ray scanner. Participants respond on the touchscreen, and the list of allowed and prohibited items grows as they continue to play.

Mitroff et al. (2015) analyzed data from the first billion trials of visual search from the game, making new discoveries about how rare targets are processed when they are presented with common foils, something that would never have been possible in the laboratory. Wolfe (1998) had previously analyzed 1 million visual search trials from across 2,500 experimental sessions which took over 10 years to collect. In contrast, Airport Scanner collects over 1 million trials each day, and the rate is increasing as the game gains popularity. In addition to answering theoretically important questions in visual attention and memory, Mitroff et al.'s example has practical implications for visual detection of rare targets in applied settings, such as radiologists searching for malignant tumors on mammograms. Furthermore, data from the game have the potential to give very detailed information about how people become expert in detection tasks.

DEVELOPING COGNITIVE THEORY BY MINING LARGE-SCALE NATURALISTIC DATA

Excerpted from *Big Data in Cognitive Science*

Intertwined Theory and Methods

Our theoretical advancements from Big Data and new methodological developments are heavily interdependent. New methodologies to answer questions from Big Data are giving us new hypotheses to test. But simultaneously, our new theoretical models are helping to focus the new Big Data methodologies. Big Data often flows in as an unstructured stream of information, and our theoretical models are needed to help tease apart the causal influence of factors, often when the data are constantly changing. Big Data analyses are not going to replace traditional laboratory experiments. It is more likely that the two will be complimentary, with the field settling on a process of recurring iteration between traditional experiments and data mining methods to progressively zero in on mechanistic accounts of cognition that explain both levels.

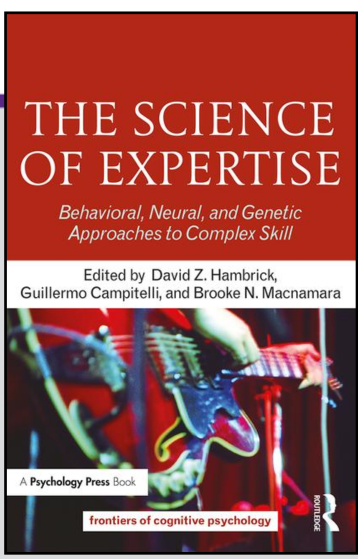
In contrast to our records from behavioral experiments, Big Data is usually unstructured, and requires sophisticated analytical methods to piece together causal effects. Digital behavior is often several steps from the cognitive mechanisms we wish to explore, and these data often confound factors that are carefully teased apart in the laboratory with experimental control (e.g. the effects of decision, response, and feedback). To infer causal flow in Big Data, cognitive science has been adopting more techniques from machine learning and network sciences. One concern that accompanies this adoption is that the bulk of current machine learning approaches to Big Data are primarily concerned with detecting and predicting patterns, but they tend not to explain why patterns exist. Our ultimate goal in cognitive science is to produce explanatory models. Predictive models certainly benefit from more data, but it is questionable whether more data helps to achieve explanatory understanding of a phenomenon more than a well-controlled laboratory experiment.

Hence, development of new methods of inquiry from Big Data based on cognitive theory is a priority area of research, and has already seen considerable progress leading to new tools. Liberman (2014) has compared the advent of such tools in this century to the inventions of the telescope and microscope in the seventeenth century. But Big Data and data mining tools on their own are of limited use for establishing explanatory theories; Picasso had famously noted the same issue about computers: "But they are useless. They can only give answers." Big Data in no way obviates the need for foundational theories based on careful laboratory experimentation. Data mining and experimentation in cognitive science will continue to be iteratively reinforcing one another, allowing us to generate and answer hypotheses at a greater resolution, and to draw conclusions at a greater scope.

CHAPTER

6

NEURAL APPROACHES IN RESEARCH ON EXPERTISE



This chapter is excerpted from

The Science of Expertise

Edited by David Z. Hambrick, Guillermo Campitelli, Brooke N. Macnamara.

NEURAL APPROACHES IN RESEARCH ON EXPERTISE

Excerpted from *Science of Expertise*

Introduction

To study the neurophysiology of expertise, neuro-imaging techniques can be very convenient tools. The aim of this chapter is to provide an overview of these methodologies used to study the neural bases of expertise and briefly present the main pattern of results found in the literature, along with their interpretation, and the associated conceptual issues. This primer chapter is specially addressed to newcomers in this domain, as we introduce what one needs to know to begin the neural study of expertise. Importantly, this is not a technical chapter; the scope is larger, as it also encompasses conceptual issues. Therefore, we only briefly present each technique and what it can achieve; but we direct the interested reader toward more technical articles or books. The chapter begins with a brief description of the main neuro-imaging techniques that have been used to study expertise and concludes with a presentation of the neural patterns of results often found in studies of expertise.

Neuro-Imaging Techniques to Scrutinize the Brain

Nowadays several techniques are used by researchers to study the cerebral activations and structures that undergird expertise. These neuro-imaging techniques can be divided into two categories: (1) those that allow studying the activations that occur in the brain—*functional techniques*—and (2) those that allow uncovering the structure of the brain—*structural techniques*.

NEURAL APPROACHES IN RESEARCH ON EXPERTISE

Functional Techniques: EEG, MEG, PET, fMRI

Among the functional techniques, *electroencephalography* (EEG) and *magnetoen-cephalography* (MEG) allow the most direct measures based on neurophysiology. Because EEG and MEG have a high temporal resolution (in terms of milliseconds) but a modest spatial resolution of 1 to 3 cm, (e.g., Siebenhühner, Lobier, Wang, Palva, & Palva, 2016), these are appropriate tools if one is interested in the time course of fairly fast processes.

EEG records the electrical activity of the brain allowing measurement of fluctuations in the electrical voltage via electrodes placed on the scalp. The electrocortical activity is measured in microvolts (μV) and then amplified by a 106 factor (e.g., Marcuse, Fields, & Yoo, 2015).

MEG also allows a direct measure of neural activity, but while EEG records the potential distribution caused by the currents on the scalp, MEG measures the magnetic fields produced by the currents' activity (e.g., Hari & Salmelin, 2012). Compared to EEG, MEG systems are expensive and cumbersome, but spatio-temporal resolution is reached without using complex head models; moreover, MEG does not require a reference because it is an absolute measure (Supek & Aine, 2014).

If one is more interested in the source of the neural signal and thus in "where" kinds of questions, *positron emission tomography* (PET) and *functional magnetic resonance imaging* (fMRI) are the tools of choice. In effect, what these techniques lack in temporal resolution (in terms of seconds), they make up for in spatial resolution (5 mm to 1 mm).

PET allows the detection of weak brain metabolic and blood flow changes (Buckner & Logan, 2001). Blood flow provides the necessary energy to neurons making it a good index of neural activity. In PET, radioactive tracers injected in the bloodstream accumulate where metabolic demand increases. Via a chain reaction involving positrons and electrons, the disintegration of the radioactive tracer causes the emission of two photons in opposite directions which induces an ionizing radiation. PET allows accurate quantification of radiotracer distribution in the brain. Then a map of brain activations (location and intensity) can be drawn. PET has a good spatial resolution, although lower than that of fMRI. The technique of fMRI indirectly measures neural activity through the so- called "BOLD" signal (Blood-Oxygen-Level Dependent). Sensory or cognitive processing is associated with local firing of neurons, which results in increased local cerebral metabolism. This requires more oxygen, so that locally the ratios between oxygenated and deoxygenated hemoglobin are modified. These changes increase the BOLD signal (Jezzard, Matthews, & Smith, 2001). In short, acquisition of the BOLD signal in response to a given task allows localization of the area of the brain involved in the task.

Structural Techniques: VBM and DTI

This second category of neuro-imaging techniques enables observation of the structure of the brain. While structural and functional techniques do not meas- ure the same substratum, the pattern of activation measured with functional techniques is thought to be linked to the structure of the brain; for the moment, however, how the linkage occurs is not clear (e.g., Wang, Dai, Gong, Zhou, & He, 2014). Moreover, concerning expertise, very few studies have investigated how practice does or does not differentially impact structure and function.

NEURAL APPROACHES IN RESEARCH ON EXPERTISE

Excerpted from *Science of Expertise*

Two types of neural structures can be observed: (1) gray matter of the brain (neurons) using *voxel-based morphometry* (VBM), or (2) white matter using *diffusion tensor imaging* (DTI).

Voxel-based Morphometry

VBM which involves *magnetic resonance imaging* (MRI), allows obtaining three-dimensional images of the brain, including gray matter, white matter, and cerebrospinal fluid. MRI relies on the magnetic proprieties of hydrogen nuclei in water molecules (present in large quantities in the brain). In the absence of a magnetic field, protons are oriented randomly, but when subjected to a strong magnetic field, they align on the same axis. The return of protons to equilibrium state engenders a radio signal that can be used to achieve detailed images of brain tissues (Brown & Semelka, 2010). VBM allows analysis of the whole brain. It compares the local density of gray matter, white matter, or cerebrospinal fluid between several groups of subjects. This process involves the spatial normalization of all images into the same stereotactic space (to create a template), the segmentation (i.e., assigning each voxel to gray matter, white matter or cerebro-spinal fluid) and the smoothing of data. Finally, statistical analyses are carried out in order to localize differences between groups.

Diffusion Tensor Imaging

This is a specific kind of modeling for *diffusion weighted imagery* (DWI), which is a variant of conventional MRI based on the tissue-water diffusion rate. In a barrier-free environment, diffusion is isotropic (the probability of displacements is the same for every direction), however it is anisotropic in white matter. Direction and amount of diffusion enables computation of a "tensor," which is an estimate of the local diffusion directions, thus leading to the so-called "DTI." Different methods are used to quantify the tensors in each point of the brain (a common method uses fractional anisotropy). A classical application of DTI is *fiber tractography* (Stieltjes *et al.*, 2001).

How to Use These Neuro-Imaging Techniques

Once one has access to these techniques, how should they be used to study expertise? The best way to understand this is to see how these techniques have been used and what kinds of patterns have been reported. As in the behavioral approach, two kinds of approach exist: (1) *cross-sectional*, where experts are compared to novices, and (2) *longitudinal*, where one measures the same individuals while they are gaining expertise. We will not go into details concerning the pros and cons of each approach, as they have already been covered in chapters on the behavioral approach. Of interest here is the specificity of the neural approach; therefore we will present two scanning paradigms used in the longitudinal approach and tease apart their respective advantages and disadvantages.

The most classic paradigm is *scanning-training-scanning* (S-T-S) (for a review of studies adopting this paradigm, see Guida, Gobet, Tardieu, & Nicolas, 2012). In S-T-S, the brain is scanned before training (*novice condition*) and after training (*practiced condition*), or additionally in between.

In contrast, some authors (e.g., Moore, Cohen, & Ranganath, 2006) prefer using a *training-scanning* (T-S) paradigm where participants are first trained concerning a task and a type of material and scanned only after training; how- ever, participants are scanned twice after training: (1) while executing the task for which they have been trained and with the same material used during the training — the *practiced condition*, and (2) while executing the task for which participants have been trained *but* with a novel kind of material for which they have no expertise — the *(pseudo)novice condition*.

Importantly, these paradigms measure the effect of practice in different ways. First, whereas studies using the S-T-S paradigm compare a physiological state before training and a physiological state after training, in the T-S studies the physiological state before practice is unknown and is not part of the experimental contrast. The two physiological states that are compared in the T-S paradigm are the (pseudo)novice condition and the practiced condition, both of which are post-training. This can be a shortcoming, because the training can also affect the (pseudo)novice condition, especially since the same task is used in the practiced condition and in the (pseudo)novice condition. Therefore, while the S-T-S paradigm seems to engender an authentic "trained vs. untrained" contrast, the T-S paradigm seems more likely to show a "trained in a task with very familiar object vs. trained in a task with less familiar object" contrast. The T-S paradigm also has interesting features compared with the S-T-S paradigm. In fact, with the S-T-S paradigm there are two confounds that do not appear in the T-S paradigm: *time and training*.

In the S-T-S paradigm, the time factor is confounded with the *practice* factor. One way of controlling this is to use a control/placebo group and to scan it twice exactly the same as the experimental group. Due to the cost of the scanning process this is rarely done. In the T-S paradigm, the time factor is not confounded, because the two scanning sessions occur at the same moment, after the training.

The second confound in the S-T-S paradigm is the training procedure once its content has been removed. Let us imagine that the training can be divided into two parts: (a) the content of the training and (b) the training without its content. In this way, an effect of training could be due to the content of that particular training (the "a" part) or just because training occurred, meaning that even an irrelevant training would cause an effect (the "b" part).

Summing up, it appears that the S-T-S paradigm has fewer methodological problems only if a control/placebo group is used. However, if one wants to avoid using controls/placebos, a solution would be to merge the two paradigms; that is, (1) scan all the participants (like in S-T-S); (2) train them (like in S-T-S); and finally (3) after the training, give them a novel material and the familiar material (used in the training), in order to scan the participants in the "untrained" situation (the novice condition, like in T-S) and scan the participants in the trained situation (the practice condition, like in S-T-S and T-S).

Now that we know more about the scanning paradigms that are used, let us turn to the changes in the brain that co-occur with the acquisition of expertise. Logically, if one compares two moments on a continuum of expertise, two changes in the brain can take place when considering one location of the brain: increase or decrease. However, when one takes into account multiple locations in the brain, then one must add *reorganization*, which can be defined as a combined pattern of increases and decreases across brain areas (Kelly & Garavan, 2005).

In the following sections, from the literature, we present and interpret the three cerebral change-patterns that are observed through functional and structural neuro-imaging techniques.

Interpreting Cerebral Patterns

Brain Decrease

Brain decrease can be functionally or structurally observed. This decrease has been reported indirectly by comparing experts vs. novices (cross-sectional studies) or directly (longitudinal studies). In the latter case, most authors use the S-T-S paradigm. No matter the techniques or comparisons, when studying the effect of expertise, the decreasing pattern is the most observed pattern, with, as we will see later, one exception—motor tasks.

Several interpretations of the decreasing pattern exist; nonetheless, there seems to be almost a consensus. As reviewed by Guida *et al.* (2012), authors tend to describe the decreasing pattern in terms of neural efficiency of the brain. This hypothesis (e.g., Buschkuehl, Jaeggi, & Jonides, 2012), which can be linked to Haier and colleagues' neural efficiency theory (Haier *et al.*, 1988) postulates that the ratio between information processing and cerebral resources can increase with practice, leading to a more efficient cortical functioning. Interestingly, Poldrack (2000) suggested that in order to speak about "neural efficiency" the brain decrease must be associated with a behavioral improvement, otherwise the decrease can be caused by other (confounded) factors rather than by efficiency.

Though this suggestion is appealing, Poldrack (2015) recently warned against the use of this concept in an explanatory fashion, as it is insufficient and necessitates a more mechanistic explanation. And, indeed, at least three cognitive explanations of brain decrease have been proposed. For Beauchamp, Dagher, Aston, and Doyon (2003) for example, the decrease of activation they detected in the orbitofrontal cortex areas suggested that skill in solving Tower-of-London problems became increasingly implicit with practice. A second way of interpreting efficiency is in terms of flexibility. Landau, Garavan, Schumacher, and D'Esposito (2007) have shown, for example, that practice impacts more unimodal and multimodal regions than primary sensory and motor regions. The first two types of regions are more flexible because they participate in top-down "modulatory and selection processes," while the latter are involved in bottom-up perceptual and motor processes, which are less susceptible to adaptation. A third way of explaining efficiency is by appealing to the concept of *chunks*. Guida *et al*. (2012) proposed that with practice there is a gradual buildup of chunks (*chunk creation*). These chunks can then be used (*chunk retrieval*) to represent and process the world with fewer cognitive and neural resources (see also Guida, Gobet, & Nicolas, 2013).

Finally, increased efficiency may be accounted for by more physiological oriented explanations such as reduction in the number of activated neurons (e.g., Garavan, Kelley, Rosen, Rao, & Stein, 2000) or by neural efficiency (e.g., Kelly & Garavan, 2005).

As noted above, decreased efficiency seems to be found in almost all the expert activities, especially at the early stages as postulated by Guida *et al*. (2012, 2013) in their two-stage framework of expertise acquisition, except specifically for motor-sensory activities where brain increase is more often observed. So let us turn to this pattern.

Brain Increase

The physiological features of increase can be considered as the reverse pattern of decrease. If one considers that activation decrease is the contraction of a neural representation, increase can be seen as an expansion of the cortical representation (Pascual-Leone, Amedi, Fregni, & Merabet, 2005). For example, in tasks such as long-term motor training that tap the primary cortices, there is a modification in horizontal connectivity (e.g., Buonomano & Merzenich, 1998) that suggests an extension of the representation. However, according to Poldrack (2000), the increase in activation can have at least one other interpretation, i.e., a strengthening of the response activation (see also Kelly, Foxe, & Garavan, 2006), and it is difficult to distinguish between the two possibilities.

No matter the technique, studies on motor expertise almost always trigger a neural pattern of increase when comparing the neuro-imaging data of experts versus that of novices. This pattern has been evident since the first study on the subject (Elbert, Pantev, Wienbruch, Rockstroh, & Taub, 1995), which showed that string-player musicians had a larger cortical finger representation of the left hand that correlated with musical expertise (starting age). Reviews by Buschkuehl *et al.* (2012) and Kelly and Garavan (2005) confirmed this pattern, showing for example that increased brain activation can be observed almost systematically in motor-sensory tasks (e.g., Ungerleider, Doyon, & Karni, 2002). Conversely, expertise in more complex tasks does not seem related to an increase of activation (Buschkuehl *et al.*, 2012; Guida *et al.*, 2012; but see Klingberg, 2010), although it is important to note that the only meta-analysis on motor tasks is not clear-cut (Yang, 2015).

Concerning structural techniques, the same pattern of increase is found. For example, in their review, Zatorre, Fields, and Johansen-Berg (2012) (see also Zatorre, Chen, & Penhune, 2007) showed that there consistently seem to be a greater gray matter volume and cortical thickness in auditory cortices of expert musicians. Bengtsson *et al.* (2005) also showed that extensive piano practice had specific effects on white matter development in the spinothalamic tract. And Zatorre *et al.* (2012) reported that generally in all the studies reviewed, the effects increased as a function of years of musical practice. This pattern was also reported by Debarnot, Sperduti, Di Rienzo, and Guillot (2014) in their review. However, given the cross-sectional nature of these studies, the link is only correlational. The same is true in various domains such as basketball (Park *et al.*, 2011), typing (Cannonieri, Bonilha, Fernandes, Cendes, & Li, 2007), or golf (Jäncke, Koeneke, Hoppe, Rominger, & Hänggi, 2009).

But more direct longitudinal evidence does exist. For example, using a longitudinal design Draganski *et al.* (2004) observed an increase of gray matter density—in mid-temporal regions and in the left posterior intra-parietal sulcus— when individuals trained at juggling for three months. A comparable result was obtained after seven days (Driemeyer, Boyke, Gaser, Büchel, & May, 2008). Interestingly, this increase can also alter white matter pathways as shown by Scholz, Klein, Behrens, and Johansen-Berg (2009) using DTI.

Zatorre *et al.* (2012) listed the candidate mechanisms for explaining these structural changes.1 For gray matter, the major candidates are *neurogenesis*, *gliogenesis*, and *synaptogenesis* changes in neuronal morphology and even vascular changes. For white matter, the major candidates are *activity-dependent axonal sprouting*, *pruning or re-routing*, and *myelination* (Fields, 2015).

NEURAL APPROACHES IN RESEARCH ON EXPERTISE

Excerpted from *Science of Expertise*

Co-occurrence of Brain Decrease and Increase

Interestingly, structural techniques have also detected more complex results, with co-occurring increase and decrease. For example, using VBM, James *et al.* (2014) compared expert musicians and novices. They reported a decreasing pattern in motor-sensory areas which co-occurred with a gray matter density increase in higher-order regions. This type of pattern is not unique; for example, when compared to individuals who do not drive taxis, taxi drivers usually exhibit greater volume in the posterior hippocampus with less volume in the anterior hippocampus (Maguire *et al.*, 2000). As noted above, this pattern of combined increase and decrease may be termed *reorganization* in studies utilizing functional techniques (Kelly & Garavan, 2005). However, in other research using structural techniques (especially in the domain of rehabilitation), reorganization is considered more synonymous with *plasticity*, and the term may be employed to designate change, which can be a simple decrease or increase.

Functional Reorganization

In this section, we present and interpret the special status of functional reorganization as a combined pattern of activation increases and decreases across brain areas (Kelly & Garavan, 2005). At present, two types of functional reorganization have been pinpointed: *scaffolding* (Petersen, van Mier, Fiez, & Raichle, 1998) and *true reorganization* (Kelly & Garavan, 2005).

The concept of scaffolding is the idea that a set of brain regions can be used by unskilled individuals in effortful performance to cope with novel task demands when attentional and control processes are needed. After a period of practice, the activation of the brain areas involved in scaffolding decreases while the activation of other brain areas may increase. Kelly and Garavan (2005) contrasted scaffolding, which fades away with practice, to true reorganization, which does not occur immediately. Two recent reviews (Guida *et al.* 2012; Neumann, Lotze, & Eickhoff, 2016) are compatible with the idea that reorganization occurs only at a late stage of practice; that is, when sufficient time has been available to practice. Guida *et al.* (2012) have tied this physiological process to the development of knowledge structures; that is, cognitive structures that allow increasing the cognitive context around incoming information, with a consequent increase in mnemonic capacities.

Relevant here is the concept of *neural context* (e.g., Bressler & McIntosh, 2007), which suggests that the context of activation around a particular brain area is also crucial. From this perspective, it is easy to understand how a brain region can have different functions if there is variation in its interaction with different brain areas. Related to expertise, this concept translates as follows: Even if an expert and a novice have exactly the same pattern of activation in one brain region, this region could be processing information in a completely different manner because its interactions with other regions could be different. In terms of functional reorganization, neural context seems therefore to be crucial, although its application still needs to be taken further.

Conclusion

This chapter has shown that no matter the kind of techniques—functional or structural—similar cerebral patterns are found in expertise acquisition. When comparing increasing and decreasing patterns, the latter is the most commonly found, especially when people start practicing (e.g., Guida *et al.*, 2012), no mat- ter the domain, with the exception of motor-sensory tasks. In this case, an increasing pattern is often observed (Buschkuehl *et al*., 2012; Kelly & Garavan, 2005). Finally, when sufficient time is available, reorganization tends to take place, and has been linked by Guida *et al*. (2012) with the development and utilization of knowledge structures. These three patterns—decrease, increase, and reorganization—that co-occur with expertise acquisition are fundamental for the understanding of neural research on expertise.