

6

How Forgetting Aids Homo Ignorans

Lael J. Schooler

Abstract

Can some functions of deliberate ignorance be achieved through processes that govern forgetting? This chapter expands on this question and considers how processes critical to encoding, retrieving, and forgetting information in memory might help to achieve some of the functions attributed to deliberate ignorance. Consideration is given to whether both deliberate ignorance and forgetting are devices that can help with “information management” (e.g., by helping with information overload). The ACT-R model of memory, which holds that human memory can be understood as an information-management system, is used to illustrate how forgetting can function as a “performance-enhancing device” (e.g., by showing how the recognition heuristic, a simple inference strategy, depends on forgetting to perform well). Constructive processes of memory, which include forgetting, are explored for their ability to regulate emotions (by putting aside or reshaping memories of past experiences) and to serve as “strategic devices” (to avoid responsibility and improve an individual’s ability to deceive more generally). Although memory shares functions with deliberate ignorance, this chapter finds that the best strategy to stay ignorant of a piece of information is to never encode it in the first place.

Introduction

Dear Dr. McGaugh,

As I sit here trying to figure out where to begin explaining why I am writing you and your colleague (LC) I just hope somehow you can help me. I am thirty-four years old and since I was eleven I have had this unbelievable ability to recall my past...

This is how AJ starts her letter to James McGaugh (Parker et al. 2006:35), as she asks for help to manage her extraordinary autobiographical memory. But why has she sought help? In her words (Parker et al. 2006:35):

I think about the past all the time....It’s like a running movie that never stops. It’s like a split screen. I’ll be talking to someone and seeing something else...Like

we're sitting here talking and I'm talking to you and in my head I'm thinking about something that happened to me in December 1982, December 17, 1982, it was a Friday, I started to work at G's (a store)....

AJ spends much of her time lost in the past, unable to resist retrieving the next memory that has been automatically cued by the last one. Her extraordinary autobiographical memory, however, lets her down in other aspects of her life. She depends on post-it notes to organize her day and forgets which of her five keys fits which lock. Nor did she do exceptionally well at school. For example, she even had trouble memorizing important dates in her history classes, although she can accurately retrieve the dates of newsworthy events from her life. She writes that “most have called it [her memory] a gift but I call it a burden” (Parker et al. 2006:35).

Technological advances mean that soon we may all have to choose whether we want extraordinary memories akin to those of AJ. This choice is central to an episode of the *Black Mirror* series (a British science fiction anthology series for television), entitled “The Entire History of You.” Set in a future where an implant behind the ear lets people replay video from their past on the back of their eyelids or on a screen for others to see, the show focuses on the problems that result from the protagonist having access to a perfect record of his life. He dwells on a parting comment made by a job interviewer and obsessively reviews scenes from a party, looking for evidence of his wife’s infidelity. In the end, he goes through great pains to remove the implant, and in so doing he deliberately chooses a degree of ignorance of his past.

Hertwig and Engel (p. 3, this volume) define *deliberate ignorance* as

the conscious individual or collective choice not to seek or use information (or knowledge; we use the terms interchangeably). We are particularly interested in situations where the marginal acquisition costs are negligible and the potential benefits potentially large, such that—from the perspective of the economics of information (Stigler 1961)—acquiring information would seem to be rational (Martinelli 2006).

Hertwig and Engel ask whether some functions of deliberate ignorance can be achieved through processes that govern forgetting? Here, I expand on this question to consider how the processes critical to encoding, retrieving, and forgetting information in memory help to achieve some of the functions Hertwig and Engel attribute to deliberate ignorance. First, consideration is given to Hertwig and Engel’s suggestion that both deliberate ignorance and forgetting are devices that can help with “information management” by helping with information overload. Thereafter, an outline is provided of the ACT-R model of memory, whose key tenet is that human memory can be understood as an information-management system (Anderson et al. 1997). Second, using the ACT-R model of memory, I illustrate how forgetting can function as a “performance-enhancing device” by showing how the recognition heuristic, a simple inference strategy, depends on forgetting to perform well. Third, I explore how

the constructive processes of memory, which include forgetting, can help with “emotion regulation” by putting aside or reshaping memories of past experiences. Fourth, these same constructive processes are then shown to work as a “strategic device” to avoid responsibility and improve an individual’s ability to deceive more generally. In conclusion, although memory shares functions with deliberate ignorance, our best strategy is to stay ignorant of a piece of information is to never encode it in the first place.

Memory Functions as an Information-Management Device

Hertwig and Engel (this volume, 2016) suggest that both deliberate ignorance and forgetting help the cognitive system with “information management.” How and why forgetting might be functional has been central to a proposal for understanding the adaptiveness of the memory system, called the *rational analysis of memory* (Anderson and Milson 1989; Schooler and Anderson 2017). This framework assumes that there is some cost, C , associated with retrieving a memory. This cost may reflect metabolic expenditure in maintaining and retrieving a memory as well as the time to search and to consider the memory. If the memory proves to be useful to the current purposes, there is some gain, G , in accessing the memory. The problem facing the memory system is to come up with a scheme that minimizes the retrieval costs while maximizing the gains. The rational analysis also proposes that the memory system can, in effect, assign some probability, P , to a memory being relevant in advance of retrieving it. Given these three quantities, an adaptive memory system would search memories in order of their expected utilities, $PG - C$, and stop considering memories when a memory with a probability P is about to be retrieved such that

$$PG < C. \quad (6.1)$$

This predicts that people will be able to retrieve most rapidly memories that are most likely to be relevant to their current needs and avoid recalling memories that are unlikely to be relevant.

It may help to communicate the idea behind the rational analysis by noting that human memory solves a problem much like the Internet search engine Google is designed to accomplish. The essential analogy being that Google searches for websites germane to the search query and the memory system searches for memories relevant in the current context. Need probability is the probability that the information contained in a particular memory (or website) is needed in a particular context. In the case of Google, context would be the collection of search terms used to initiate search, perhaps along with the documents, emails, and other materials Google uses to tailor search results. In the case of the cognitive system, context would be salient elements of the environment or memories active in working memory. The memory system prioritizes

the most relevant memories by making them available in working memory, by making others more or less available in long-term memory, and perhaps by forgetting others completely. Guided by the constraints of the rational analysis, Anderson developed a memory model for the ACT-R theory of cognition, a unified framework for simulating human cognition and behavior (for an overview, see Anderson et al. 1997). Here I present a stripped-down version of this memory model, but one that retains enough of the spirit to the original to investigate the parallels between forgetting and deliberate ignorance.

If i refers to a memory, then its activation, A_i , reflects the probability of the memory being needed in a given context. The key equation for calculating A_i , the activation of memory i , is

$$A_i = B_i + \sum_{q \in C} W_q S_{qi}, \quad (6.2)$$

where B_i is its base level activation, reflecting the past frequency and recency of needing the memory; S_{qi} is the strength of association between q (an element of the current context, C) and memory i ; and W_q reflects the attention paid to element q in the context. B_i can be thought of as memory i 's resting activation level:

$$B_i = L \ln(F_i) - D \ln(R_i). \quad (6.3)$$

Essentially, a memory's resting activation is an increasing function of F_i , how frequently the memory has been resident in working memory, either through being encoded from the environment or retrieved from long-term memory, and a decreasing function of R_i , how long it has been since the memory was last resident in working memory. The parameters L and D can be thought of as learning and decay rates, respectively. Equation 6.3 shows that all things being equal, a memory will be increasingly accessible the more time it has spent in working memory, and subsequently less accessible the longer it has been since it was last in working memory.

The second part of Equation 6.2 corresponds to the association of various elements of the context to the memory:

$$S_{qi} = \frac{p(i|q)}{p(i)}. \quad (6.4)$$

Association is captured by associative ratios, $p(i|q)/p(i)$. The denominator of this ratio, $p(i)$, is the base rate probability of needing memory i across all past contexts. The numerator is the conditional probability, $p(i|q)$, of needing memory i in the presence of some cue, q . The larger the ratio, the better the indicator q is that memory i will be needed. The overall associative strength of the context to the memory is the sum of all the associative ratios of each of the individual cues in the context.

The probability of retrieving a memory increases with increasing activation, whereas the time it takes to retrieve a memory decreases with increasing activation;

that is, higher activation values correspond to faster retrieval times. In short, understanding forgetting as a form of “information management” was central to the development of the ACT-R model of memory. ACT-R lends quantitative precision to how this information management works for human memory.

Forgetting Functions as a Performance-Enhancing Device

A properly functioning memory achieves many of the functions that Hertwig and Engel attribute to deliberate ignorance (this volume, 2016). For example, Schooler and Hertwig (2005) have shown how ACT-R’s model of memory can be a “performance-enhancing device.” Schooler and Hertwig (2005) explored whether forgetting could aid a simple decision strategy known as the recognition heuristic (Goldstein and Gigerenzer 2002). In its simplest form, the recognition heuristic is a strategy to rank two alternatives according to some criterion of interest. The classic example is ranking two cities according to their population. Suppose you are asked to choose which city, Durban or Johannesburg, has more inhabitants. Personally, I could only guess. In contrast, if I were asked the same question for Cape Town or Tembisa, I would wager Cape Town has the greater population based on the observation that I recognize Cape Town but not Tembisa. This strategy of relying on recognition works to the extent that recognition is correlated with the criterion of interest, which in this case is population.

Whether there is a strong correlation between the criterion and recognition depends on a chain of correlations: larger cities tend to be mentioned in the media and conversations more often than smaller cities, and how often a city is mentioned correlates with how likely it is to be recognized. To illustrate how this chain of correlations works, we can take the number of results that Google retrieves when queried with the name of a city as a proxy for how often it is mentioned. For these South African cities, there is a perfect correlation between population and the number of results retrieved: Cape Town (127 million results, 3.4 million inhabitants), Durban (90 million results, 3.1 million inhabitants), and Tembisa (4.2 million results, 500 thousand inhabitants).¹ With the correlation of environmental frequency and population in place, let us turn to recognition. Equation 6.3 implies a positive correlation between activation and environmental frequency, but a negative correlation between activation and recency. We can model whether or not a city is recognized by whether or not the activation of a memory associated with the name of that city exceeds the retrieval threshold described in Equation 6.1. Suppose that our actual exposure to these city names in a given year is one millionth of the results that Google returns; that is, 127 mentions for Cape Town, 90 mentions for Durban, and 4 mentions for Tembisa. According to Equation 6.3, if we ignore forgetting

¹ See <https://www.geonames.org/ZA/largest-cities-in-south-africa.html> (accessed January 27, 2019).

(i.e., setting D to 0) and set the learning rate, L , to the convenient value of 1, the respective activations for Cape Town, Durban, and Tembisa would be 4.8 ($= \ln 127$), 4.5 ($= \ln 90$) and 1.4 ($= \ln 4$). For convenience, assume a recognition threshold of 1.5, so that Cape Town and Durban would be recognized, because their activations are above 1.5, but Tembisa would not be, because its activation falls below 1.5. In this case, the recognition heuristic would correctly decide that Cape Town and Durban are larger than Tembisa, but would have to guess when deciding between Cape Town and Durban. Next, assume another year goes by with the same environmental frequencies, so that the total number of times the three cities are mentioned in the course of two years is 254, 180, and 8, raising their activations to 5.5, 5.2, and 2.1. Focusing on frequency alone, the recognition heuristic would have to guess when deciding among all the cities, because now Tembisa's activation of 2.1 would exceed the recognition threshold of 1.5. However, when forgetting is taken into account, as captured by the second part of Equation 6.3, the base level activation decays over time. Consider the activation associated with Tembisa on the first day of the third year. Assuming a decay rate of .2 and a recency of 90 days since the last mention of Tembisa (i.e., $90 \approx 365/4$): on the first day of the third year, the activation of Tembisa $1.2 (= \ln 8 - .2 \times \ln 90)$, would fall below the recognition threshold, enabling the recognition heuristic to be applied. The heuristic would now correctly choose which city has the most inhabitants when Tembisa is one of the options. In sum, forgetting helps maintain the correlations required for the recognition heuristic to be an effective decision strategy by reducing the chances that some memories (i.e., those associated with smaller cities) will be retrieved. As we will see next, reducing the chances of retrieving memories can help regulate emotions.

Memory Functions as an Emotion-Regulation Device

With respect to affect regulation, Hertwig and Engel (this volume, 2016) point to examples of people deliberately wanting to be ignorant of medical test results, because knowing the results could force them to face emotionally taxing decisions. In another example, they suggest that people may not want to know the pay of their colleagues so that they can avoid envy. These external sources of emotionally disturbing information would at least require making an appointment to take a test or actively searching for a colleague's pay. In contrast, the "marginal acquisition costs" of information stored in memory are close to 0. How then can a person protect themselves from emotionally disturbing information lurking in their memory?

Within the ACT-R framework, an interaction between the strategic behaviors under conscious control (at least to some extent) and the mechanisms inherent to the architecture could help control the retrieval of distressing memories. Again, it helps to think about the process of memory retrieval as a

Google search. The conscious mechanisms would be akin to making sure that searches do not include terms that are likely to return upsetting websites. For example, I might avoid searching with terms such as “Fox News” or “ISIS.” Google may also have mechanisms, of which I am unaware, designed to protect me from upsetting information by forgetting the most disturbing information that the Internet has to offer.

To illustrate how ACT-R’s memory processes could lead to forgetting that could help guard against the retrieval of emotionally disturbing memories, consider an anecdote from my past, including some specific details that we will return to later in the chapter. It is an event I would prefer to forget, has some emotional content, but is not all that embarrassing, at least given the passage of time. No doubt you can dredge up an embarrassing memory of your own. The situation involved my twelfth grade English class and its recitation of part of Chaucer’s *Canterbury Tales* for a school assembly. Mr. McCown, our teacher, wanted to have musical accompaniment. Having for many years played the flute, but never in public, I volunteered. At the rehearsal, the first notes out of my flute sounded much like a child’s first attempts to make a sound by blowing across the top of a Coke bottle. To the delight of my friends sitting high in the gymnasium bleachers, my performance went downhill from there, ending with Mr. McCown saying, “I’m not here to give music lessons.” I took the sheet music with me, practiced it under the direction of my flute teacher, and was redeemed the day of the school assembly with a passable performance. Although now, knowing what I do about how memory works, I have my doubts that the assembly turned out as well as I remember that it did. I will refer to this as the flute memory, although within ACT-R it would be modeled by a complex of associated memories.

Let’s first consider the implications of Equation 6.1, the stopping rule $PG < C$, where C is the cost associated with retrieving a memory, P is the probability of needing a memory, and G is the gain associated with retrieving the memory. Note that in the ACT-R framework there is a one-to-one mapping between a memory’s activation and P . For most ACT-R analyses, G has been assumed to be fixed and the same for all memories. However, consider what happens when this assumption is relaxed. My flute memory is definitely tinged with the embarrassment of that day. These negative emotions could be seen as reducing the gains, G , associated with retrieving the memory. The reduction in G means that increased levels of activation would be needed to retrieve my flute memory in comparison to a more positive memory. In Norby’s (2015) review of functional forgetting, he mentions that memories with negative emotions tend to be retrieved somewhat less well than those associated with positive emotions. Now, imagine a memory with strong negative emotional content, such as a memory of an assault. These negative memories could take G from being a net gain to a loss. In this case, the expected value of the memory could be negative, implying that the memory would be unlikely to be retrieved however low the retrieval cost, C , might

be. Curiously, when G is negative, the higher the activation, the less likely it would be that the corresponding memory would be retrieved. The combination of a negative G and the standard ACT-R retrieval mechanisms would be one way to implement repression in the ACT-R framework. I want to emphasize that ACT-R does not make any hard predictions here. For example, the system could ascribe positive gains to a horrific memory or the memory could be rehearsed to such an extent that even the smallest gains would put its activation above threshold. As this is to the best of my knowledge the first time that manipulating G has been proposed as a way to incorporate inhibitory mechanisms into ACT-R, it is an open question as to whether G should be under deliberate control.

Decidedly, deliberate strategies that could be used to avoid negative memories operate through the part of Equation 6.2 that handles the association of a memory to the current context. In ACT-R, my flute memory would be represented as a collection of memories, associated with each other and various retrieval cues, analogous to the key words used in a Google search. The cues might be Mr. McCown, a flute, gymnasium bleachers, Chaucer's *Canterbury Tales*, or feeling embarrassed. A deliberate strategy would be to avoid cues associated with the flute memory (e.g., by avoiding going back to my high school gymnasium). An interference-based strategy would be to reinforce associations between these cues and other memories. For example, when I think of Mr. McCown, I could try to remember all the positive memories I have of him as a teacher (he was among the best) or the burgundy shirt and jeans he routinely wore to class. In ACT-R, such an interference-based strategy would increase the associations between Mr. McCown and these other positive memories and simultaneously decrease the associations that connect Mr. McCown and the flute memory, lowering the activation of the memories that compose the flute memory, and, in turn, lowering the probability that they would be retrieved. The base level strength would reinforce the effects of these deliberate strategies. Each time a memory is retrieved, it increases its base level activation. Preventing these retrievals would allow the base level activation of these memories to be forgotten (i.e., decay away), perhaps even leading to completely forgetting the disturbing memory.

Another strategy to avoid disturbing memories depends on the observation that memories are as much constructed as they are retrieved. There is ample evidence that people often mistake an imagined event for a real one and, further, that these false memories may be facilitated by techniques encouraged in some forms of psychotherapy (Loftus 1997) and legal investigations (Ceci and Huffman 1997). In the flute example, I can think of two places where I suspect these constructive processes may be operating. The first is that when I initially recounted the memory for this chapter, I did not remember what Mr. McCown was wearing. Now, as I think back on the day, I clearly remember Mr. McCown in his burgundy outfit. Most likely, I filled in the missing information about what he was wearing with my general knowledge about what he tended to

wear, and most likely he was dressed all in burgundy that day. As I mentioned earlier, I now have a clear memory of my performance at the assembly going okay. Does it really matter whether it did go well or that I just imagined that it did? The constructive nature of memory provides us with the opportunity to fill in forgotten information or replace remembered information with new information that is less emotionally charged.

A related idea is being used to treat patients with posttraumatic stress disorder. The basic technique is to have patients remember their traumatic stay on an intensive care unit and then give them propranolol, which causes a reduction in stress hormones, so that when the memory is reconsolidated it becomes dissociated with the high levels of stress that were originally associated with it (Gardner and Griffiths 2014). A careful understanding of the constructive nature of memory could guide the development of a toolbox of techniques that may help us deliberately forget the mundane memories of a flubbed high school performance and even the traumatic memories of being in an intensive care unit. Next, we will see how some of these same constructive memory processes and techniques can help us strategically avoid responsibility by improving our ability to deceive others by deceiving ourselves.

Memory Functions as a Strategic Device to Avoid Responsibility and Enhance Deception

Hertwig and Engel (this volume, 2016) suggest that deliberate ignorance can be used as “a strategy for *avoiding liability* in a social or even legal sense.” Forgetting can serve a similar function. For example, in his senate confirmation hearings to become the attorney general of the United States of America in 2017, Jeffrey Sessions answered “I don’t recall” or something to that effect on 47 distinct topics (Lapowsky 2017). Whether he had truly forgotten the information or not, the answer reduces his accountability and risk of exposure by preventing subsequent questions on the topic. In the intelligence committee confirmation hearings, Sessions failed to recall a conversation with Sergey Kislyak, the Russian ambassador to the United States at the time. Undoubtedly, the committee would have wanted to learn more about the details of the conversation that could have shed light on connections between Donald Trump’s close associates and the Russian government. We will never know for sure whether Sessions forgot his conversation with Kislyak or not. Even if Sessions one day reveals that he did remember the conversation, it could well be that he remembers a reconstructed memory of the event. Whereas feigning forgetting may be a deliberate strategy used to avoid liability, true forgetting may serve the important evolutionary function of deception. Trivers (2000, 2011a) has argued on evolutionary grounds that deception is an important capacity that helps conspecifics acquire resources from each other. On one side is an individual trying to deceive and on the

other an individual trying to detect the deception. The cycle of deception and detection leads to an evolutionary arms race. As argued by von Hippel and Trivers (2011), some of the cues used to detect deception result from the cognitive demanding compensations required to mask signs of nervousness. For example, when the deceiver attempts to control signs of nervousness in their face, the pitch of their voice may also rise. They argue that because many of the signs of deception are mitigated through self-deception, evolution would favor individuals who can deceive themselves. Self-deception reduces the cognitive load of deception, not only by reducing the cognitive demands associated with masking nervousness, but also by the demands required to consciously maintain both the false narrative behind a deception and the true narrative. Although there were a mere 47 events that Sessions could not recall, there were 87 occasions on which he was asked about these 47 events.³ If he were pretending not to remember, he would be faced with the cognitively demanding task of keeping track of how he responded previously. One way to reduce cognitive load is to maintain a single narrative of past events, supported by memories consistent with the deception. According to von Hippel and Trivers (2011), the constructive nature of memory, as described above, is one of the primary mechanisms used to massage memories into a story that supports deception. Essentially, if the deceiver believes their story to be true then they are less likely to produce the telltale signs associated with deception, such as those associated with nervousness and an inconsistent retelling of past events.

Conclusion

Memory processes, including forgetting, can achieve functions that have been ascribed to deliberate ignorance. Each of the functions—information management, performance enhancement, emotional regulation, strategic deception—illustrate the potential benefits of limiting the amount of information an individual considers. The chapter began with examples of people with near perfect memories, implicitly contrasting their memories with the run-of-the-mill memories of regular folks. We often complain about how much we forget, losing sight of how phenomenally good we are at retrieving the information we have stored in our memories. In point of fact, most of us are blessed and cursed with extraordinary memories. The extraordinary memories that most of us possess mean we need to carefully guard the information we allow ourselves to know about. Notwithstanding the various ways in which we can reconstruct some memories and forget others, once information is in memory there is a good chance that it will not be forgotten and that it will be retrieved effortlessly. Once we know, for instance, that we have tested positive for the HIV virus or the Tay–Sachs gene, we are unlikely to be able to forget the test results. Despite the fact that memory and forgetting share functions with

deliberate ignorance, our memories are so good that we need to use the tools of deliberate ignorance to shield our memories from learning information we want to remain hidden.

