8
Making Ourselves Useful

8.1 Beliefs & Reality

The theory of perceptual reality monitoring (PRM), introduced in Chapter 7, connects two important notions of consciousness: subjective experience, and our rational grasp of reality. The former is what we have focused on so far. But the latter is no less relevant. Studies in the social sciences, psychiatry, and clinical psychology have focused on this notion of consciousness for over a century. To many, that's where the concept of consciousness really comes from, or at least that's what makes the topic so interesting. A theory of consciousness would do well to allow something meaningful to be said about this connection between the two notions.

The local theorist may complain that I'm changing the topic. We set out to understand consciousness in order to say something about the Hard Problem. There, subjective experience is what matters. Rationality does not. But I'm not saying that consciousness is rationality. PRM is about subjective experience, as promised. Nor am I saying that consciousness is what you believe. Rather, it is the mechanism by which you *potentially* form beliefs about what you perceive. These beliefs are rational in the sense that they *feel* justified; it seems to make sense to believe them and to act according to them, at least from the subject's point of view. This is how consciousness is connected to rationality.

The goal of this chapter is to flesh out this connection in more detail. A better understanding of this connection can inform clinical and social applications. It also makes clear why PRM is a "centrist" theory (Chapter 6), that is, a happy medium between extreme options. It accounts for subjective experience, without writing off broader issues like the role of consciousness in cognition and rational behavior. I will also argue that it is only with this more extended version of PRM that we can fully account for the experiences of emotions and volition.

There is an alternative to this line of thinking, which is to say: consciousness is just subjective experience. For those who use the word to refer to anything related to rationality, they are just using the word *entirely* differently. It is true that some words refer to totally different things (e.g., *palm* can refer to

In Consciousness We Trust. Hakwan Lau, Oxford University Press. © Hakwan Lau 2022.
DOI: 10.1093/oso/9780198856771.003.0009

a kind of tree or a part of our hand). But are the notions of "consciousness" in these different contexts really entirely unrelated in such a way? At times many contemporary neuroscientists working on consciousness come close to taking this extreme view. But remember from Chapter 1 Section 1.5, that definitions are never hard and fast matters. There is a cost to defining ourselves into a corner, making ourselves irrelevant to our academic neighbors. As we will see, there is just no need to do so.

8.2 Symbolic Causal Narratives

Rational agents act according to their beliefs and desires. For example, if we desire an apple, and we believe that there is an apple in front of us, we may proceed to grab the apple. But other beliefs are also relevant: For example, is it polite given the social context? How much money would it cost? Could the apples be poisonous?

Some of the relevant beliefs are not perceptual beliefs. They don't concern what we see or otherwise feel at the moment. They may even concern things that do not exist in the concrete, such as social etiquette, money, and the fear of death by poison. And yet they no doubt impact our rational actions and decisions.

Nor are these more abstract beliefs entirely unrelated to our perceptual beliefs. The two sets of beliefs have to somehow hang together in a *coherent way*. If we believe that it is socially inappropriate for someone to grab the apple in the present context, but we see another person doing it, either we doubt the social norm we believe in, or we doubt what we're seeing correctly—or we form the judgment that the person grabbing the apple must be rude. Somehow, we form an overall *story* that is more or less coherent. Things we believe in have to make sense together. They have to roughly "add up."

We can call this coherent web of beliefs our *narrative* of reality. In narratives, things are expressed in causal terms. When we say that President Obama got fired up in a rally and started leading the crowd to cheer, we mean that his enthusiasm *caused* him to act that way, and his action *caused* the crowd to follow. To believe that something is socially inappropriate is to believe that *if* certain norms are violated, there will *then* be certain undesirable consequences. To believe that something is poisonous is to believe that consuming it will *cause* sickness or death.

But causal models are notoriously difficult to build from observations alone. Just because X precedes Y doesn't mean that X causes Y at all; they could be both caused by something else, with the effect on Y just being more delayed.

Or they could be mere coincidences. It would help if we could manipulate X at will. And our actions do impact our world. But as far as the beliefs we just talked about are concerned, we can't move President Obama's level of enthusiasm up and down in order to assess the consequences accordingly. We can't turn a social norm on and off at will to see if people's behavior changes. Sometimes these are stories just told to us. They might have already happened, and there's no room for intervention.

My former colleague at UCLA Judea Pearl has spelled out the logic for building these kinds of models for causal inferences (Pearl and Mackenzie 2018). These models are crucial for counterfactual reasoning. That is an extremely important ingredient for intelligence. To make good decisions, we need more than a grasp of the known facts. We need to be able to think about hypothetical scenarios: For example, *what happens if I try this new solution? What happens if we try something else instead?*

The computational methods proposed by Pearl are elegant, but they also highlight how demanding they are, in terms of the amount of data needed, *if* we are to build the models from the ground up. So, instead, many assumptions need to be made. The role of data is primarily to arbitrate between plausible alternatives. These alternatives can be expressed schematically in some "graphs" (Figure 8.1).

This means that to empirically inform causal reasoning, we probably need to conceptualize the relevant events at a *symbolic* level. Our sensory representations are analog and detailed. For the narrative models to work, we would do well to include only the key facts, and exclude uncertain, noisy information. Some abstraction and simplification are probably needed to summarize things for causal reasoning.

Importantly, the sensory representations driven by imagination and perception may be similar. But at the causal reasoning level they make a world

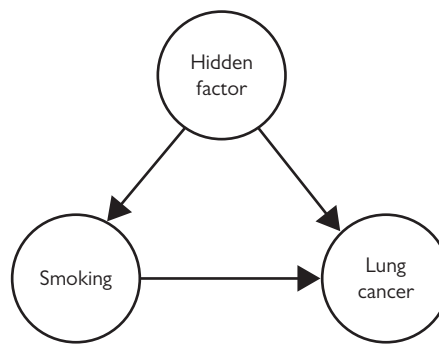

Figure 8.1 An example "graph" representation of what may cause lung cancer

of difference. Some hard and fast decisions need to be made about what the sensory activity really represents before the information is passed on to the narrative level.

This may be how the perceptual reality monitor described in Chapter 7 contributes to our narratives of reality, by boiling down rich perceptual information into something worth believing and thinking about further. In this sense, consciousness may be the "gating" mechanism through which perception impinges on higher cognition.

8.3 Split-Brain Patients & Confabulations

I have not given proof that this is exactly how we represent reality in the brain. In Section 8.2 we went through some engineering considerations as to what may be a plausible general architecture, based on our current knowledge in the computer and cognitive sciences. Ideally, we shall investigate the neurocognitive mechanisms more directly in the future. Some existing work is relevant and congruent (reviewed in LeDoux and Lau 2020). But perhaps just as telling are some classic patient studies.

In split-brain patients with their corpus callosum severed, information could be selectively presented to just one of the hemispheres. When information is presented to the right hemisphere, patients tend not to report being aware of it, because language functions mostly reside on the left hemisphere. However, such information can evidently have an impact on behavior. In one famous study, the patient was asked to pick a picture that was most relevant to what was shown (Volz and Gazzaniga 2017). A snow scene was shown to the (nonconscious) right hemisphere, and a chicken claw was shown to the left hemisphere. The patient picked a shovel. That made sense given the snow scene. But because the patient was not aware of seeing the snow scene, he reported that the shovel was relevant to the chicken claw, because it could be used to clean up the chicken excrements!

This kind of confabulation behavior has been found in other patients too, such as those who suffered from hemispatial neglect. In one study a patient was presented with two houses, the left of which was on fire in one (Marshall and Halligan 1988). Because of the neurological condition, the patient reported that both houses looked similar. However, when asked which house the patient would prefer to live in, the house on fire was avoided. And yet, the patient could not really give a convincing answer to justify the choice.

One interpretation is that the coherence-seeking narrative system can only take into account what one is conscious of (Liu and Lau 2021). When

nonconscious information impacts behavior, it creates a difficult situation for the system to make sense of the situation. Hence the confabulatory justifications are somewhat funny. In this sense, our narrative system is often not fully rational, just *quasi*-rational.

I cannot say that these examples provide decisive evidence for a narrative system exactly as described in Section 8.2. But my point here is more modest: something akin to a coherence-seeking narrative system probably exists in some form. To fully understand consciousness, we would do well to learn more about how such a system may interact with subjective experiences.

8.4 Narratives as "Consciousness"

Why should we care about these narratives? One reason is that there is a long tradition in the social and mental health sciences of identifying this narrative mechanism as "consciousness" itself.

When Marxist theorists talk about "false consciousness," it is not about sensory hallucinations (Lukacs 1972). It is about how certain people fail to comprehend reality correctly, and to act rationally according to how resources and power are actually allocated. They may see the same physical objects and hear the same sounds as the ruling class. They see the same movement of goods, food, and money. But they misunderstand the underlying causes and effects. They underestimate the far-reaching consequences *if* they were to collectively stop working for the ruling class. Their consciousness is "false" because they got the causal narratives wrong.

Likewise, when the great sociologist Émile Durkheim wrote about "collective consciousness," it referred to the common narratives of reality shared between members of a group (2014). It was not about the science fiction-like possibility that different people may merge together into a single conscious entity, as if they could share a singular stream of subjective experiences.

Importantly, according to these views our narratives are functionally relevant. They impact our rational decision-making. Getting the narratives wrong is detrimental to our bargaining behavior. It causes us to be exploited. It can make social cooperation difficult.

In the psychiatry and clinical psychology tradition, "consciousness" is also often understood in terms of these narratives. In particular, psychoanalysts claim that there is an "unconscious" mind that is opaque to direct cognitive access and control (Alexander 1948). But the "conscious" mind is not equated with only ongoing sensory experiences either. Whatever one can access through direct introspection, including abstract knowledge, are considered

"conscious." So, like the political theorists, they also often use the term *consciousness* to refer to our rational (or rationalized) narratives. But psychoanalysts suggest that the "unconscious" mind is just as capable of sophisticated forms of reasoning, as if it also has its own independent narrative system of some sort.

Nevertheless, psychoanalysis has not become an empirically successful science. This leaves us doubting if there really are such things as fully fledged "nonconscious" narrative mechanisms. The view I propose is that there isn't.

However, on a more practical level, these two questions are also left open: *are there "nonconscious" ways to change our "conscious" narratives?* And how about the opposite: *can our "conscious" narratives impact "nonconscious" processes too?* Below I will argue positively for both: our "conscious" narratives and at least some "nonconscious" processes influence each other. This is true especially for affective processes, which often seem to happen outside of our conscious control. This may in a sense suggest that the psychoanalysts aren't entirely wrong. Not only does this matter for our theory of consciousness, but also these questions have clinical and societal implications too.

Because of this tricky terminological issue of the different usages in different fields, I will put "conscious" and "nonconscious" in quotations, as I did above, when I refer to this notion of (quasi-)rational narratives that guide deliberate actions, rather than subjective experiences (our default usage of the term *consciousness*).

8.5 Self, Actions, & Responsibility

One of the most common and stubborn misunderstandings against PRM, and its higher-order relatives, is that we overintellectualize consciousness. The charge is that we conflate simple subjective experiences with explicit introspection about oneself (Malach 2011). But the criticism is misplaced because the theory does not actually assume anything like that (Lau and Rosenthal 2011). The self-monitoring mechanism hypothesized in Chapter 7 is automatic and implicit (Sections 7.8 and 7.9). That is, some mechanism in your prefrontal cortex needs to know the dynamics of your sensory activity in order to make a perceptual decision. But you *as a person* do not need to make an effort to do so. Conscious experiences just happen to you, even if you don't try to think about them.

But how, then, do we account for our explicit sense of self-awareness? Certainly, people can introspect and explicitly think about themselves too. The narrative system proposed here can perhaps accommodate this because

one of its main functions is to deal with the challenging problem of causal reasoning. We think of ourselves as causal agents. We believe that our actions and decisions have causal consequences. In Chapter 5 we pointed out that empirically this is actually a somewhat open-ended question; our sense of agency may well be partially reconstructed after the facts. Nevertheless, regardless of whether our conscious intentions are causal, we no doubt believe that they are. One plausible view is that this comes from our higher-level narratives about ourselves, rather than directly from the motor control system. Truthfully or otherwise, we model ourselves *as* a causal agent in our self-narrative (Laurie Paul, Tomer Ullman, Julian Freitas, Josh Tenanbaum, personal communication).

This is another reason why understanding this narrative system is so important. Philosophical theories often link our sense of free will and moral responsibility to consciousness. But it is unclear if subjective experience per se is the relevant notion (Levy 2014). Instead, our overall sense of agency may arise at the narrative level. This may be why it is relatively malleable (Chapter 5 Section 5.3). And yet, it does not mean that our sense of agency is necessarily always constructed after the facts, as a mere illusion. That is because the narrative system can have causal impacts on both our subjective experiences and rational behavior. As such, disturbances in this sense of agency can have devastating effects, as, for example, observed in psychosis.

8.6 Schizophrenia the Really Hard Problem

Schizophrenia is a severe mental illness that affects as many as close to 1% of the global population (Saha et al. 2005; Schultz, North, and Shields 2007). Many people think of it as *the* disorder of consciousness, probably because of the symptoms of psychosis. We already addressed the possible mechanisms for sensory hallucination in Sections 7.8 and 7.9. But psychosis can also involve the loss of sense of agency, as well as various forms of delusions at a higher-cognitive level. Why do these symptoms often occur together in the same patients?

Chris Frith argued that one underlying mechanism for psychosis may be that patients fail to accurately anticipate the consequences of their self-generated (mental) actions (1992). Accordingly, the sensory consequences become unexpectedly salient and, therefore, sometimes are mistakenly attributed to being caused by external forces. This elegantly accounts for the variety of symptoms, from hallucinations to loss of agency, and the delusory belief that thoughts have been "inserted" into one's minds. In all cases, the idea is that the

patients themselves actually generated the imagery, action, or thought. But their brains failed to recognize that fact.

Our account here is similar. But here we factor the symptoms into two levels: the sensory and the cognitive (narrative). As we have discussed in Sections 7.8 and 7.9, these two levels may depend on partially overlapping mechanisms in the prefrontal cortex. Hallucinations, which are mostly auditory in schizophrenia, have early sensory correlates. But they are probably also caused, at least in part, by failures of the perceptual reality monitor. The "discriminator," as described in Chapter 7, mistakes noise or inner speech for external voices. This accounts for the symptoms at the sensory level.

At the narrative level of higher-cognitive reasoning, I suspect a wider variety of more complex problems may arise in psychosis. As mentioned in Section 8.2, the role of the reality monitors may be to select and simplify information for causal reasoning. When they fail to provide reliable information for this computationally demanding process, different sorts of catastrophic errors can occur. In terms of signal processing, we can say that for causal narrative reasoning, noise is *multiplicative* rather than additive. That is to say, on this analytic and symbolic level, we can't always average out the noise like we do in early sensory processing. Overall, a certain degree of coherence may still be maintained, even when the inputs are noisy, leading to confabulatory responses. But the errors on this level may become so unpredictable and bizarre that, ultimately, they may impact one's general ability to think and behave rationally.

This way, we can more easily explain why hallucinations do not always co-occur with psychosis. But one problem is that this account is much less parsimonious than Frith's elegant model. But indeed, the point here is exactly that more effort in understanding the narrative mechanism is needed. Fortunately, there is a relatively familiar way for us to intuitively appreciate and perhaps even manipulate the operations at this causal narrative level: through our language and culture.

The medical anthropologist Tanya Luhrmann has described that, in different cultures, certain individuals apparently experience auditory "hallucinations," even though they are never formally diagnosed as such (Luhrmann et al. 2015). Instead, many of them relish in their ability to "hear voices," which they consider an important part of their societal role and religious practice. In many cases, this difference in cultural beliefs allows these individuals to enjoy productive and rewarding lives without the stigma of "madness."

Perhaps these reports suggest that the sensory and cognitive components of schizophrenia are indeed somewhat independent; for these individuals, the "hallucinations" do not simply go away, even though the problems are

somewhat resolved at the "narrative" level. Arguably, one could say that these individuals still suffer from undiagnosed "delusions." But this runs the risk of unfairly imposing our own views on otherwise well-functioning people, who do not consider themselves to be suffering.

Can these thought-provoking findings translate into treatments in the clinic? Unfortunately, the effectiveness of current psychotherapy for psychosis is somewhat limited. Antipsychotic drugs targeting prefrontal dopaminergic functions remain the standard prescription. For interventions at the psychological or narrative level, there have been some exciting recent developments (Craig et al. 2018), but we await to see robust effects replicated in large-scale studies. It is also possible that our narratives are so embedded within our culture that changing them would take more than the patient's brief interaction with a therapist—at least for the moment, until we can understand the narrative mechanisms better.

Meanwhile, there may be another venue where psychology may help. If our view is right, even if there is some degree of independence between the sensory and cognitive levels, the former may be more primary; subjective experiences are associated with the gating mechanisms for downstream narrative processing. As such, disturbances at the narrative level are perhaps typically caused by faulty inputs from the sensory level. Because the reality monitoring and perceptual metacognitive mechanisms are relatively better understood, one can develop simple psychophysical tasks to test for their functioning in both health and disease (Hoven et al. 2019). Along these lines, for example, Koller and Cannon (2021) have recently found that paranoid individuals indeed showed specific deficits in metacognitive processing in recognition memory tasks. Given present knowledge, we do not expect to be able to fix these mechanisms with precision whenever they break down. But patients with schizophrenia often first go through what is called a "prodromal" stage, before the symptoms become fully fledged (Larson, Walker, and Compton 2010). Early detection with such psychophysical tasks can facilitate preventive treatment.

8.7 How Affective Experiences & Narratives Interact

The above discussion suggests that social stigma may be one reason why psychotherapy alone may not resolve all the conflicts at the narrative level. In modern societies, unfortunately, people often react to psychotic behavior with strong negative emotions, such as distress or fear. This is so

despite the fact that, in the absence of substance abuse problems, psychosis alone isn't strongly associated with violent behavior (Fazel et al. 2009). Given this cultural bias, it may be difficult from the patient's point of view to see things as going alright.

Specifically, this breaks down into three key points: i) emotions influence narrative processing; ii) narratives influence emotions; and iii) emotions can be contagious. That is to say, for a patient suffering from psychosis in modern societies, it may be difficult to form a very positive self-narrative because one usually goes through significant emotional distress. Some of this negative affect may come from other people, who do not see the situation in a very positive light. But all the same, the resulting emotion may also influence the patient's own feelings toward their condition. Let us unpack these points.

The first point may be straightforward. The narratives we've been discussing are often self-narratives. When we feel afraid, typically the only immediate belief that we can form is just that we are afraid; we may not know why or what really causes our fear. The causal analysis may take place at least in part at the narrative level. In seeking coherence, we tend not to accept easily that we are frightened for no reason. Instead, the mind seeks plausible explanations.

The social psychologist Jonathan Haidt has given the analogy that our "conscious" decision-making is a bit like riding an elephant (2012). As "rational" agents, we'd like to think we are in control of our emotions. But the slow and powerful animal we are riding has its own mind. We can nudge it toward our desired directions here and there. But in the end, we are often just going along for the ride. We make up post hoc rationalizations to justify our "decisions." But maybe emotions are really the driving forces.

This idea traces back to David Hume, and has a huge influence on the social sciences, maybe in particular in political theory (Gauthier 1979). Its wisdom rings just as true today: much of our political and moral debates do come down to affect, often more than we realize. If people feel a certain way about certain political issues, they may contrive to come up with the arguments to support how they feel. This probably happens more often than we are prepared to admit. In this sense, the logical arguments are secondary, and it is emotion that is truly fundamental to, for example, social change or elections.

But as Pizarro and Bloom (2003) have pointed out, the influence can go the other way too. Narratives at the cognitive level can also influence emotions (i.e., narratives influence emotions, the second point above). Our emotional reactions to the same set of sensory events often depend on how we see the overall narrative. If someone says something rude and aggressive to us,

our reactions can be anything from anger, fear, guilt, or remorse, to disinterested sadness. It all depends on how we interpret the situation, the history of what happened, and what we expect may happen next. It may involve counterfactual reasoning too, e.g., *What if I tell the person to calm down? How likely would it work? Or how about I just run?* Our preliminary answers to questions like these seem to determine how we feel. If such implicit thinking does take place, however imperfect it may be, it must happen pretty fast.

That is to say, emotions are not just simple "gut reactions"; we should not confuse simple physiological reactions with full-blown emotions. When we are physiologically aroused, we are sometimes just as likely to be scared or falling in love. To find out which is the case, our cognitive understanding of the context matters (Schachter and Singer 1962).

Finally, it is well-known that emotions are contagious (Hatfield, Cacioppo, and Rapson 1993; the third point above). Some may think this reflects mostly automatic mimicry. When people look afraid, we may sometimes feel scared too, without thinking. But perhaps there is also some rationale to contagion, even at the narrative level. Let's say you visit a new country and you find that people are generally mortified by the sight of squirrels. Instead, they seem to find rats cute. Perhaps it makes sense for you to draw the inference: maybe squirrels are poisonous there, and rats are relatively hygienic and harmless? Maybe eventually you will—and should—learn to be afraid of those (hypothetically) deadly squirrels too.

Despite the "quasi-rational" nature of these ways of interactions, the logic isn't always so transparent to us. Therefore, we may overestimate how "conscious" our narrative system really is. Our narratives are often influenced by emotions in ways rather opaque to us, beyond our control. Although we have the ability to regulate our emotions at will to some extent, its reach is not unlimited (Gross 2002).

8.8 Affective Learning, Homophily, & Culture

The limits of our emotional insights and regulation may be related to one of BF Skinner's arguments for behaviorism. Late in his career, Skinner speculated about "consciousness," and even openly explored Freudian ideas (Overskeid 2007). But prior to that, in a debate against the role of conscious thoughts in psychology, he mentioned that the behaviorist's point was not that these high-level constructs are unobservable and, therefore, cannot be studied (Blanshard and Skinner 1967). Rather, the worry was that they may not provide as much leverage for systematically predicting and controlling behavior.

In contrast, since Pavlov, psychology has made tremendous progress on our understanding of the mechanisms underlying the learning of simple emotions like fear (Rescorla 1988). In general, these mechanisms follow the principles of associative conditioning. When a neutral stimulus is consistently paired with something frightening, we learn to be afraid of the neutral stimulus too.

Of particular interest is one form of conditioning called vicarious learning (Olsson, Knapska, and Lindström 2020), which we have already described in our squirrel and rats example in Section 8.7. The idea is that we may learn to react with certain emotions by observing how others react to the same stimuli or situations.

Because of the ways narratives and emotions interact, through vicarious learning, different people can learn to synchronize their narratives too. This is somewhat turning the problem on its head. But I suspect it is actually a rather common phenomenon. For example, when I was only 9 years old, I saw the news on TV about a peaceful protest being cracked down on by an authoritarian regime, leading to a massacre. No doubt that back then, I had not thought through the concepts of justice, liberty, power, and mercy. But I distinctly remember learning then that anger and sadness were the "appropriate" social reactions. If someone reacted to the same tragic news with elation and joy, or sheer indifference, that person would no doubt be shunned by the people I love and respect. Without a word spoken, this already helped strangers ensure that they share a similar narrative of what really happened.

Of course, in society, we can talk through our narratives too. We can debate about them, and make sure we're all on the same page. But the full narratives are often complex, and our communication skills may be limited. On the other hand, the vicarious learning of emotions is based on simple and robust mechanisms (Olsson, Knapska, and Lindström 2020). Additionally, there is the factor of homophily. That is, we tend to sort ourselves into groups of similar people. We hang out with people we find agreeable. As I just mentioned in the example in the last paragraph, our emotional responses probably play a significant role in helping us identify like-minded people too.

These may be some of the ways through which our cognitive processes are so powerfully modulated by culture (Heyes 1993). Because of the mechanisms of associative learning, the very simple acts of sharing a meal or watching a movie together may have more impact on our "consciousness" than we may intuitively expect. Linguistic communication is no doubt important. But shared subjective experiences also contribute greatly to how common narratives are formed.

8.9 "Consciousness" Fast Versus Slow?

So, we have gone through some key concepts, regarding how "consciousness" (narratives) and consciousness (subjective affective experience) may interact. It may be useful to think of how this relates to similar models too.

For example, the Nobel laureate Daniel Kahneman distinguished between a fast (System 1) and a slow (System 2) thinking system (2011). The former is relatively intuitive, automatic, and emotion-like. The latter is more rational and analytical. The narrative system we've discussed so far may map well to Kahneman's System 2.

Likewise, in the literature on the computational models of learning, there is also a distinction between model-free and model-based learning (Dayan and Berridge 2014). In model-free learning, the causal relations between events aren't explicitly represented. The learning is more akin to statistical associations, driven by the principles Pavlovian (and Skinnerian) conditioning described in Section 8.8. Model-based learning is more similar to the narrative system described so far.

One caveat is that these views are typically not so strongly committed to the nature of subjective experience. Subjective experience may or may not accompany System 1 (i.e., fast, intuitive) reasoning; nonconscious representations (in the sense of lacking subjective experience) and unreflected conscious percepts are both likely governed by the same simple model-free associative learning principles. But if a perceptual event is nonconscious, according to PRM it will probably not be selected (by the perceptual reality monitor) for making downstream impact on System 2 or complex, model-based learning.

Another difference between these views and PRM is that some may see Systems 1 and 2 as working in parallel. Likewise, model-based and model-free learning systems may also operate side-by-side, in some form of competition or cooperation. On PRM though, the narrative system is a late-stage process downstream in a hierarchical architecture, relative to early perceptual processes. That is, it takes selected early perceptual information as input. However, this late-stage process probably also feeds back to early sensory processing (LeDoux and Lau 2020). So, simple perceptual experiences may show some level of coherence too.

We already argued for this kind of feedback modulation in the case of emotions earlier in Section 8.7. In the case of perceptual experiences, one anecdotal consideration is that dreams also tend to be somewhat structured; the subjective experiences in dreams do not seem like random sensory "noises." One explanation could be that the coherence comes from the narrative system.

If that's true even during sleep, perhaps subjective experiences can never be completely understood in isolation from the narrative mechanisms too.

The last point may be one reason why we should reject behaviorism, despite Skinner's argument described in the Section 8.8. To change psychological behavior, it may be useful to target the associative learning mechanisms for affective responses, which are better understood. However, there are occasions where the narratives themselves play important causal roles too, or they may be the very targets we want to ultimately change. In Sections 8.10-8.12, we will quickly review some examples as to how we can apply these concepts in clinical and societal contexts.

8.10 Fear & Trauma

Let us start with cases in which it is desirable to reduce the intensity of some emotional experiences. People who went through life-threatening traumas may experience "flashbacks" of the incidents, which can trigger very unpleasant and intense experiences of fear. Sometimes, it is as if the trauma is being "relived." But these are not the only possible symptoms. Posttraumatic stress disorder (PTSD) can also impact a person's cognition, including one's self-image and autobiographical memories (Sutherland and Bryant 2008). In severe cases, it can lead to suicidal thoughts and behavior. As such, psychotherapy focusing on the cognitive or narrative levels is evidently helpful in some cases (Kar 2011).

That said, a common treatment is (some variant of) exposure therapy, which we discussed briefly in Sections 5.12 and 5.13. The idea comes from simple Pavlovian conditioning principles. If a trauma-related cue (e.g., a specific weapon) is presented repeatedly without harm, one may "unlearn" previously associated fear with the cue. Essentially, the focus is on prenarrative-level sensory processing. In Sections 5.12 and 5.13, we specifically described a way to achieve this nonconsciously, using the method of decoded neurofeedback (DecNef).

Why may this kind of treatment be effective? Theoretically, it has been suggested that the traumatic events may be encoded into two kinds of memory representations: sensory and conceptual (Brewin, Dalgleish, and Joseph 1996). In PTSD the sensory representations may dominate, which accounts for the relatively uncontrolled nature of the memory process. When the memories are involuntarily "recalled," they seem intrusive and "relived" as if they are presently occurring. They seem to be disconnected with narrative-level contextual processing. Or they may be so dominant

they "hijack" normal cognitive processing. If that is correct, it may be advantageous to selectively target the abnormally dominant early sensory associative mechanisms.

In this theoretical context, there are predictions we can test too. If conscious and nonconscious processes operate entirely in parallel, targeting nonconscious sensory mechanisms may never change the conscious subjective experiences (Taschereau-Dumouchel, Liu, and Lau 2018). As we mentioned in Section 5.13, indeed currently we do not yet know if methods such as DecNef can affect more than physiological responses. However, if conscious and nonconscious mechanisms are common at the early sensory level, we may expect DecNef to eventually be able to change subjective experience too. From there, perhaps even narrative-level processing could be impacted. So, this is an empirically open-ended question that we may be able to address in the near future.

8.11 Placebo Pain & the "Crisis of Neurology"

Contrast PTSD with chronic pain. The latter may seem a lot less "psychological." But pain is also subject to placebo manipulations (Price, Finniss, and Benedetti 2008). That is, the sheer idea that they are being treated can paradoxically reduce pain, even in the absence of a real treatment. This effect is in fact very robust and common, which is why modern medicine generally adopts double-blinded procedures; in establishing the effectiveness of a new treatment, we need to rule out these powerful placebo effects. Using neuroimaging, it has been shown that the placebo effect of pain relief was correlated with reduced activity in early somatosensory regions too (Atlas and Wager 2012). So, narrative-level cognitive processes seem to influence pain experience.

The blurriness of the boundary between the psychological and early sensorimotor processing is highlighted in another class of disorders that are sometimes labeled by neurologists as "functional" or "psychogenic" (Edwards, Stone, and Lang 2014). For example, in functional movement disorders a patient may make unwanted and apparently involuntary tics. However, anatomically there are no identifiable causes. Standard physiological measures like the electroencephalogram may also detect no problems. Intriguingly, in some cases when the patients are distracted, the tics may go away. This seems to suggest that the issue is "psychiatric" or "psychological." But, in fact, these patients are typically perfectly lucid and seem to have no trouble making rational decisions.

There has been some evidence that functional movement disorders are linked to early childhood trauma (Kranick et al. 2011). But certainly most cases of traumatic experiences do not lead to such selective and unique impact on the motor system. The mechanisms are currently unclear. And yet functional movement disorders are extremely common, affecting up to 30% of all patients seeking help for movement disorders (Hallett 2019). Our limited perspective on this widespread clinical phenomenon has led some authors to describe this as a "crisis of neurology." There's an urgent need to better understand how high-level psychological and cognitive factors influence basic neurological symptoms.

8.12 Economics & Political Polarization

Earlier, in Section 8.7, I mentioned the ideas of Haidt and Hume, according to which much of our moral and political reasoning may be driven by intuitions and emotions. But I also pointed out that narratives may also influence our emotions.

In some ways, in politics, often it is the narrative that ultimately matters. My former colleague at UCLA Davide Panagia calls it *narratocracy* (2009); in a sense, society is governed by the stories we share. Take racism as an example. Some people may have a gut feeling that it is dangerous to be in a neighborhood populated by people of a different race. Statistically, perhaps this is actually true in some cases. The rate of violent crime may be higher there. In this sense, our "nonconscious" associative learning mechanisms might have picked up nothing but factual information. But the important questions to ask are: What causes the higher crime rate in the first place? Have those people of the said different race been given the same opportunities? Have they been treated fairly? These questions involve causal reasoning and can perhaps only be resolved at the narrative level.

The Nobel Prize–winning economist Robert Shiller (2017) also recognizes the power of narratives in driving economic events. He likened many major changes in the financial market to pandemics. Narratives can indeed go viral, leading to mass fear or optimism, which, ultimately, can have major impacts on the stock market and beyond. Interestingly, they also fade away in time, as would be predicted by epidemiological models.

But of course, figuring out these narratives are at times messy; especially because we cannot do experiments on historical events. In the end, we may just sort ourselves into our very own echo chambers of choice, which gives us a sense of belonging but are none closer to the truth. Another former colleague

at UCLA Jared Diamond (2019) has opined that the breakdown of face-to-face communication may be one factor in why modern politics is so polarized, especially in the United States. Different groups simply no longer see the same reality; their narratives are no longer in sync.

As I suggested earlier (Section 8.8), the simple act of sharing a meal or watching a movie together may well help us sync up our "consciousness" in ways we don't intuitively expect. Nevertheless … if only things were this simple. Our current understanding of this putative narrative system is limited. How narratives can affect society, and how they interact with consciousness, remain active areas of research within the social sciences (Clough and Halley 2007; Clough 2008; Hoggett and Thompson 2012). Recent work on using a cognitive neuroscience approach to understand the relationship between metacognition and political polarization seems especially promising (Rollwage et al. 2019).

8.13 Chapter Ending Remarks: Freud, Marx, and Panpsychic Qualia

This chapter covered a lot of ground, but, admittedly, I was only able to do it superficially. It may seem like I'm going against the very premise of this book, which is to argue for an empirically grounded account of consciousness. Much of what is said in this chapter is speculative.

But the purpose here is exactly to bring out this fact: so much is at stake, and yet we know so little.

Historically, the notion of "consciousness"—as in the sense of our rational grasp of reality—has not received a lack of attention. Rather, many great scholars have written insightfully on the topic. Unfortunately, Marxist ideas stimulated real-world "experiments" that did not turn out well; the so-called communist revolutions have led to atrocities still felt today. In the case of the Freudians, it is unclear if empirical truthfulness was ever a priority (Eysenck 1991; Crews 2017).

One important lesson, though, is that Freud's writing remains just as eloquent and stimulating today, and continues to impact the arts and humanities, as well as popular culture. All the same, the sheer lack of empirical rigor alone has attracted tremendous backlash, stifling mental health research for decades (Eysenck 1991). Great ideas can sometimes cause great harm.

Accordingly, many rigorous scientists choose to avoid the topic altogether. It may be correct to point out that one ought not to throw the baby out with the bathwater. But guilt by association is a common phenomenon.

In the introduction, we discussed the "revival" in the 1990s. But is our "new" science of consciousness speaking to these historically important issues in the social and clinical sciences? Are we restoring the balance or merely shifting the attention further away?

Meanwhile, "consciousness" continues to be discussed, with or without our empirical input. In clinical assessment of mental illnesses, self-reports are often treated as supplementary information, giving way to more "rigorous" and "objective" biomarker approaches (forthcoming review w/ Ledoux et al). But these biomarkers are not always in concordance with self-reported subjective experiences (Taschereau-Dumouchel, Kawato, and Lau 2020). And yet, one of our primary goals should be to make the subjects *feel* better. Besides reducing excessive physiological responses, we also want their pain and fear to go away *subjectively*. Our lack of meaningful engagement in these areas means that ultimately the patients are the ones who suffer.

Ironically, those who favor scientifically untestable notions of "pure qualia" over "consciousness" do not really sidestep the issues. By defining consciousness a priori in nonfunctional terms, they cut themselves off from any chance of being able to make meaningful connections to the social and clinical sciences. By not keeping their empirical tallies straight, they also risk making the very same mistakes that ultimately brought down the Freudian empire.

My hope is to convince you that we can avoid this "double fault." The theory of perceptual reality monitoring (PRM) can be related to "consciousness" as we traditionally understand it in other disciplines. Much work still needs to be done. But here I have sketched out how these connections can at least be made and further studied in principle; subjective experience and our (quasi-) rational narratives causally interact with each other, in systematic ways. It may therefore be advantageous to understand both in conjunction, rather than in isolation. This is especially so if we intend to manipulate either people's subjective experiences or their narratives; sometimes it may be most effective to change one by influencing the other, or we may try to target both at the same time for potential synergistic effects. For many basic psychological phenomena, mental disorders, and societal problems, our understanding will ever remain incomplete if we do not also consider the narrative mechanisms.

Given the potential utility of these applications, it would be unwise for us to define our discipline into isolated obscurity. Fans of panpsychism or other forms of nonfunctional "qualia" may challenge that we focus too much on high-level cognition, betraying our "roots." Is PRM capable of answering their philosophical concerns too? We will address this in Chapter 9, our next and final chapter.

References

Alexander F. *Fundamentals of Psychoanalysis*. W.W Norton, 1948.

Atlas LY, Wager TD. How expectations shape pain. *Neurosci Lett* 2012;**520**:140–148.

Blanshard B, Skinner BF. The problem of consciousness: A debate. *Philos Phenomenol Res* 1967;**27**:317–337.

Brewin CR, Dalgleish T, Joseph S. A dual representation theory of posttraumatic stress disorder. *Psychol Rev* 1996;**103**:670–686.

Clough PT. The affective turn: Political economy, biomedia and bodies. *Theory, Culture & Society* 2008;**25**:1–22.

Clough PT, Halley J. *The Affective Turn: Theorizing the Social*. Duke University Press, 2007.

Craig TK, Rus-Calafell M, Ward T et al. AVATAR therapy for auditory verbal hallucinations in people with psychosis: A single-blind, randomised controlled trial. *Lancet Psychiatry* 2018;**5**:31–40.

Crews F. *Freud: The Making of an Illusion*. Profile Books, 2017.

Dayan P, Berridge KC. Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. *Cogn Affect Behav Neurosci* 2014;**14**:473–492.

Diamond J. *Upheaval: Turning Points for Nations in Crisis*. Little Brown, 2019.

Durkheim E. *The Division of Labor in Society*. Simon & Schuster, 2014.

Edwards MJ, Stone J, Lang AE. From psychogenic movement disorder to functional movement disorder: It's time to change the name. *Mov Disord* 2014;**29**:849–852.

Eysenck HJ. *Decline and Fall of the Freudian Empire*. Transaction Publishers, 1991.

Fazel S, Gulati G, Linsell L et al. Schizophrenia and violence: Systematic review and meta-analysis. *PLOS Med* 2009;**6**:e1000120.

Frith CD. *The Cognitive Neuropsychology of Schizophrenia*. Psychology Press, 1992.

Gauthier D. David Hume, contractarian. *Philos Rev* 1979;**88**:3–38.

Gross JJ. Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology* 2002;**39**:281–291.

Haidt J. *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. Knopf Doubleday Publishing Group, 2012.

Hallett M. Functional movement disorders: Is the crisis resolved? *Mov Disord* 2019;**34**:971–974.

Hatfield E, Cacioppo JT, Rapson RL. Emotional contagion. *Curr Dir Psychol Sci* 1993;2:96–100.

Heyes CM. Imitation, culture and cognition. *Anim Behav* 1993;**46**:999–1010.

Hoggett P, Thompson S. *Politics and the Emotions: The Affective Turn in Contemporary Political Studies*. Bloomsbury Publishing, 2012.

Hoven M, Lebreton M, Engelmann JB et al. Abnormalities of confidence in psychiatry: An overview and future perspectives. *Transl Psychiatry* 2019;**9**:268.

Kahneman D. *Thinking, Fast and Slow*. Macmillan, 2011.

Kar N. Cognitive behavioral therapy for the treatment of post-traumatic stress disorder: A review. *Neuropsychiatr Dis Treat* 2011;**7**:167–181.

Koller WN, Cannon TD. Paranoia is associated with impaired novelty detection and overconfidence in recognition memory judgments. *J Abnorm Psychol* 2021. 130;3:273–285 https://doi.org/10.1037/abn0000664.

Kranick S, Ekanayake V, Martinez V et al. Psychopathology and psychogenic movement disorders. *Mov Disord* 2011;**26**:1844–1850.

Larson MK, Walker EF, Compton MT. Early signs, diagnosis and therapeutics of the prodromal phase of schizophrenia and related psychotic disorders. *Expert Rev Neurother* 2010;**10**:1347–1359.

Lau H, Rosenthal D. The higher-order view does not require consciously self-directed introspection: Response to Malach. *Trends Cogn Sci* 2011;**15**:508–509.

LeDoux JE, Lau H. Seeing consciousness through the lens of memory. *Curr Biol* 2020;**30**:R1018–R1022.

Levy N. *Consciousness and Moral Responsibility*. Oxford University Press, 2014.

Liu K, Lau H. A framework for understanding agency. In: PW Kalivas, MP Paulus, JR Lupp (eds), *Intrusive Thinking: From Molecules to Free Will.* Vol 30, pp. 229–245. MIT Press, 2021.

Luhrmann TM, Padmavati R, Tharoor H et al. Hearing voices in different cultures: A social kindling hypothesis. *Top Cogn Sci* 2015;**7**:646–663.

Lukacs G. *History and Class Consciousness: Studies in Marxist Dialectics*. MIT Press, 1972.

Malach R. Conscious perception and the frontal lobes: Comment on Lau and Rosenthal. *Trends Cogn Sci* 2011;15:507; author reply 508–509.

Marshall JC, Halligan PW. Blindsight and insight in visuo-spatial neglect. *Nature* 1988;**336**:766–767.

Olsson A, Knapska E, Lindström B. The neural and computational systems of social learning. *Nat Rev Neurosci* 2020;**21**:197–212.

Overskeid G. Looking for Skinner and finding Freud. *Am Psychol* 2007;**62**:590–595.

Panagia D. *The Political Life of Sensation*. Duke University Press, 2009.

Pearl J, Mackenzie D. *The Book of Why: The New Science of Cause and Effect*. Basic Books, 2018.

Pizarro DA, Bloom P. The intelligence of the moral intuitions: comment on Haidt (2001). *Psychol Rev 2003*;**110**:193–196; discussion 197–198.

Price DD, Finniss DG, Benedetti F. A comprehensive review of the placebo effect: Recent advances and current thought. *Annu Rev Psychol* 2008;**59**:565–590.

Rescorla RA. Behavioral studies of Pavlovian conditioning. *Annu Rev Neurosci* 1988;**11**:329–352.

Rollwage M, Zmigrod L, de-Wit L et al. What underlies political polarization? A manifesto for computational political psychology. *Trends Cogn Sci* 2019;**23**:820–822.

Saha S, Chant D, Welham J et al. A systematic review of the prevalence of schizophrenia. *PLOS Med* 2005;**2**:e141.

Schachter S, Singer JE. Cognitive, social, and physiological determinants of emotional state. *Psychol Rev* 1962;**69**:379–399.

Schultz SH, North SW, Shields CG. *Schizophrenia: A Review.* 2007. American family physician. 75(12):1821–1829.

Shiller RJ. Narrative economics. *Am Econ Rev* 2017;**107**:967–1004.

Sutherland K, Bryant RA. Autobiographical memory and the self-memory system in posttraumatic stress disorder. *J Anxiety Disord* 2008;**22**:555–560.

Taschereau-Dumouchel V, Kawato M, Lau H. Multivoxel pattern analysis reveals dissociations between subjective fear and its physiological correlates. *Mol Psychiatry* 2020;**25**:2342–2354.

Taschereau-Dumouchel V, Liu K-Y, Lau H. Unconscious psychological treatments for physiological survival circuits. *Curr Opin Behav Sci* 2018;**24**:62–68.

Vincent Taschereau-Dumouchel, Matthias Michel, Hakwan Lau, Stefan Hofmann, Joseph LeDoux (in press). Putting the 'Mental' Back in 'Mental Disorders': A Perspective from Research on Fear and Anxiety. Molecular Psychiatry.

Volz LJ, Gazzaniga MS. Interaction in isolation: 50 years of insights from split-brain research. *Brain* 2017;**140**:2051–2060.