Pablo Fernández Velasco[1]

Attention in the Predictive Processing Framework and the Phenomenology of Zen Meditation

Abstract: *In this paper I will use the phenomenology of Zen meditation (zazen) to look at the role of attention within the predictive processing (PP) framework. Section 1 introduces PP, according to which the brain is a dynamical, hierarchical, hypothesis-testing mechanism. Section 2 discusses the current proposal that attention is the process of precision optimization (Hohwy, 2012) and presents some of the challenges for this theory. Section 3 introduces zazen and uses some of the emerging patterns of its phenomenology to clarify the workings of attention, with a special emphasis on the difficulty of maintaining a relaxed and homogeneous state of attention. I claim that this difficulty corresponds to a hyperprior that leads to the expectation of a given level of uncertainty in the world, which in turn pulls attention towards distracting input. Section 4 looks at research about cognitive control and meditation, and concludes that the agent can attempt to impose a global strategy (such as a globally distributed precision expectation) that can affect the assignment of precision expectations, but that this*

Correspondence:
Email: p.fernandezvelasco@gmail.com

[1] This work was completed whilst at School of Philosophy, Psychology and Language Sciences, University of Edinburgh

Journal of Consciousness Studies, **24**, No. 11–12, 2017, pp. 71–93

assignment ultimately depends on a complex interplay of different factors. Section 5 discusses some possible challenges for the claims of this paper and Section 6 is a conclusion followed by possible future directions.

1. Introduction to Predictive Processing

I prefer to work in cafés. I find the silence of libraries oppressive, and there are too many distractions at home. There is a particular café in Edinburgh (Affogato, an ice-cream parlour in the West End) where I often go to do work. The music is low and easy to ignore, the place has comfortable couches, and the coffee is good. There is a designer, Laura, who is in there quite often, so I always say hello when I see her. Two weeks ago I went in on a particularly sunny day (it was during an unprecedented row of five sunny days), and Laura was sitting in her favourite spot. I asked her how her last job was going (she had to design a website for a shoe company).

'Excuse me?' she answered, confused.

I looked again and started: Laura wasn't Laura. On the couch there was a woman who looked remarkably like Laura: same height and same narrow oval face, but now that I looked closer, slightly different hair colour. After a moment I realized that she didn't even look that remarkably like Laura. A bit embarrassed, I excused myself and sat down to work.

That was a strange occurrence, but nothing extraordinary. A similar thing happens when you drink what you thought was a latte but turns out to be tea with milk, or when you hear your phone ring, take it out of your pocket, and see that no one is calling you. The tea even tastes like coffee for a second (to me at least), and then comes the surprise. In this paper I will use a theoretical framework for how the mind works that might (among many other things) explain these kinds of events. This theoretical framework is often referred to as predictive processing, and it claims that what we experience as the world is determined by our own predictions of how we expect the world to be, or, to be more precise, that '*conscious perception* is determined by the prediction or hypothesis with the highest overall posterior probability' (Hohwy, 2012, p. 3). In this section I will sketch out some of the characteristics of this framework, to then focus (in the following section) on how it can help us explain the role of attention.

Predictive processing (PP) is an ambitious framework that aims to explain not only perception, but also other aspects of the human mind such as emotions, or the production of action (which is accomplished by proprioceptive prediction; see Friston, 2010). There is a lot of support for PP coming from different fields, but it is not in the scope of this paper to defend the whole PP framework. Instead, the aim of this paper is to use the phenomenology of Zen meditation to clarify how attention fits within the PP framework.

According to PP, the brain is a dynamical, hierarchical, hypothesis-testing mechanism. There is a lot of uncertainty and noise in the world that surrounds us. However, there are patterns in the uncertainty, which makes prediction possible. If PP is right, the brain uses previous information to make predictions about the world, which go top-down and sideways (i.e. from the most abstract areas of the mind down to sensory organs, and in parallel within a level), and get updated with the error of those predictions, which go sideways and bottom-up. The aim of the system is to reduce prediction error (by changing the predictions or by sampling the world to make it coincide with the predictions).

In the earlier example, my generative model predicted that Laura was sitting on a couch, which is what I experienced. This prediction was incorrect, so my senses sent prediction errors up the system, which refuted the hypothesis that Laura was sitting on the couch, provoking a revision of the hypothesis and a new prediction (namely, that someone other than Laura was sitting on the couch). Of course, hypotheses don't just get updated with starts and sudden surprises. Predictions are continually updated with error, and most times our phenomenal experience is smooth. If the person on the couch had not said anything, I would still have eventually realized that it wasn't Laura, only the process wouldn't have been so sudden. However, in some exceptional situations, the errors are not strong enough (or the assumptions behind the predictions are too strong) to change the predictions. An example of this is the hollow-mask illusion. Take a cheap plastic mask (the kind that some people wear at Halloween when they don't want to work hard to make their own costume) and look at the back of it. Most people who do this perceive the face (which is concave) as convex. A likely explanation using the PP framework is that there is a deep-rooted hyperprior ('an expectation about the world that is stable, and often at a high degree of abstraction', Wilkinson, 2014) that makes us expect faces to be convex (i.e. the nose sticks out instead of caving into the face), which impedes the error from changing the hypothesis (Clark, 2015).

We humans do not just perceive the world as if through a screen. We are of the world and in the world that we perceive. In the PP model, action occurs when the system predicts the sensory (including proprioceptive) consequences of the said action, so that the agent needs to act to minimize prediction error. Back to my favourite café, my system might make a prediction that I will reach out for my coffee and drink it. To avoid this prediction from being mistaken (therefore generating prediction error), I will reach out for my coffee and drink it. Of course, this might not be the introspective sensation of the process at a personal level, but the process at a subpersonal level (often unconscious) that is responsible for the action. This process happens simultaneously at multiple levels: the prediction 'I am drinking coffee' has many constituent predictions around arm movement, hand to cup affordance, drinking motion, and, ultimately, the taste of coffee in one's mouth. Interestingly, action is also a way of sampling the world to revise or confirm predictions. Often, a prediction initiates an action (me drinking coffee) that in turn confirms an hypothesis about the world (that the drink inside the cup is coffee, and not, say, tea with milk). Considering both action and perception as part of the process of maximizing Bayesian model evidence (which in information theory is known as 'active inference') 'separates the problems of optimizing action and perception by assuming that action fulfills predictions based on inferred states of the world' (Friston *et al.*, 2016).

2. Consciousness and Attention

Attention and consciousness are difficult to define, and the relationship between them hard to pin down. In what follows, I will use Hohwy's definition of both terms within the framework of PP as a starting point (Hohwy, 2012).

The brain, if PP is correct, is constantly comparing different hypotheses, and what we perceive consciously is the most probable hypothesis given the current waves of sensory evidence and what we already knew about the world. To determine the probability of a hypothesis, it is not enough to know about the error of said hypothesis, but also about the 'precision' of that error. The reliability of prediction error signals varies according to task and context, including ambient environmental conditions. Thus, the different errors of a prediction don't always have the same variance and are not weighted equally for updating predictions. The inverse of 'variance' is called 'precision', and the more precise prediction error is assessed to be, the more likely it is that it will cause a change in the organism's hypothesis. Of course, 'precision' cannot be known with certainty, so what will be used to weight prediction error is 'precision expectation'. The error that arises from those sensory units that are expected to give very precise errors is weighted more than the error from sensory units that are expected to give imprecise errors. If an agent is half-deaf but has very good sight, for example, the error arising from her auditory organs is going to have less weight than the error arising from her vision when it comes to updating the system's hypothesis.

Following this reasoning, an agent would not just try to minimize prediction error, but also to optimize 'precision'. One way to do this would be to selectively sample sensory input that is expected to have high 'precision'. Hohwy defines attention as the process of precision optimization. Attention (precision optimization) is then closely related to conscious perception (the overall most probable hypothesis), which captures 'the common sense notion that conscious perception and attention are intertwined and also the notion that they are separate mechanisms' (*ibid.*, p. 4).

This notion might be easier to grasp if we think about it in the following way: the things that we pay attention to play a bigger role in shaping our conscious experience than the ones that we don't pay attention to. In the case of exogenous (bottom-up) attention, stronger signals are expected to be more precise. For that reason, signals with a large contrast (either temporal or spatial) grab our attention. We pay more attention to a painting on an empty wall (spatial contrast) than to the blank wall, and to a person entering an empty room (temporal contrast) than to the unchanging floor.

Hohwy argues that endogenous (top-down) attention 'works as an increase in baseline activity of neuronal units encoding beliefs about precision' (*ibid.*, p. 7). The belief that precise input will appear in a given area of the visual field will mean that once that input appears it will be given high weight. At the personal level, more attention will be paid to that given area, and the expected input from that given area will preferentially populate consciousness. We saw before that the process of predicting an action and fulfilling it involves many constituent predictions (in the case of drinking coffee these were arm movement, cup–hand affordances, taste, etc.). In a similar way, the process of precision optimization works simultaneously at various levels that include not only neural gain, but also the mechanics of eye gaze, foveation, and fixation.

In a famous 1975 experiment, Neisser and Becklen showed participants two overlapping videos. One was a close-up of a hand game and the other was a distant ball game. It was fairly easy for the participants to pay attention to one of the films while ignoring the other and to switch attention between films, but it was difficult for them to attend to both films at the same time (Neisser and Becklen, 1975). This seems consistent with the PP account of attention. There are two competing hypotheses: the hand game and the ball game. In PP terms, when a participant paid attention to one of the scenes, that scene was given higher precision expectation, and populated consciousness, while the error remaining from the other scene was dismissed as imprecise. When a participant switched attention between scenes, the precision expectation assignment changed, making the other scene populate consciousness. The reason why it is hard to follow both scenes at the same time is that the hypothesis that in the same spot there is a close-up hand game and a far-away ball game going on is ecologically impossible.

The Neisser and Becklen case is reminiscent of binocular rivalry (the phenomenon in which visual perception alternates between different images presented to each eye). In an epistemological review of binocular rivalry, Hohwy, Roepstoff and Friston aim to explain the phenomenon with the framework of PP. They claim that binocular rivalry occurs when:

> (i) there is no single model or hypothesis about the causes in the environment that enjoys both high likelihood and high prior probability and (ii) when one stimulus dominates, the bottom-up, driving signal for that stimulus is explained away while, crucially, the bottom-up signal for the suppressed stimulus is not, and remains as an unexplained but explainable prediction error signal. This induces instability in perceptual dynamics that can give rise to perceptual transitions or alternations during rivalry. (Hohwy, Roepstorff and Friston, 2008, p. 1)

Discussing the Neisser and Becklen case, Ransom, Fazelpour and Mole (2016) argue that, in principle, the signal from the unattended scene is as precise as that of the attended scene. Here, it is important to clarify that the decisive variable is not precision, but *precision expectation*. Just as the prediction of an action gives rise to the action, to change the participant's attention it is enough for the participant (at a subpersonal, unconscious level) to predict higher precision in the unattended film. This higher prediction responds to the context. It might be that the participant 'decides' (at a higher level) that the hand game has become more interesting, and then the error coming from the hand game scene becomes more reliable. It is, of course, more reliable now to predict the hand game scene. This clarification doesn't fully cast out the problem. It still seems arbitrary to say that one film has higher expected precision than the other. If such subtle fluctuations in precision expectations can push for an attention switch, how can attention ever stabilize on one of the scenes?

A possibility is that the assignment of expected precision is self-reinforcing. The attended signal becomes more precise because it is attended to. It is not unlike using a magnifying glass: imagine struggling to read the small writing of a medication leaflet. The whole leaflet looks equally 'imprecise', just squiggles. Then you decide to read the warnings section using the magnifying glass. That section becomes clear (precise), while the others fade even more into the background. Likewise, expectations of imprecision should have the opposite result. This will be important later on, as there is something here that (I shall argue) the current PP story does not adequately accommodate. For the moment, the moral is just that expectations of high precision correspond to increases in gain of the prediction error signal from the attended sensory channel. This is plausible, and Hohwy might agree with this theory, but it still leaves the residual problem that self-fulfilling prophecies of the sort presented by Ransom, Fazelpour and Mole would prevent any attention switching between scenes, unless there is another top-down mechanism that decides how to allocate precision expectation, which (they argue) would be external to the PP account. It also leaves open the question of what the mechanism that decides to attend to a given signal, and that thus kick-starts the self-fulfilling loop, is (Ransom, Fazelpour and Mole, 2016).

Hohwy doesn't answer this problem directly. He suggests that it can 'be disadvantageous for a system to be stuck in active inference and neglecting to revisit the bound on surprise [prediction error] by updating the model' (Hohwy, 2012, p. 9), but he doesn't elaborate on the idea much further. It is true that being stuck in a self-fulfilling loop could be disadvantageous for a system (if, for example, the participant were stuck looking at the hand game scene and a hungry tiger entered the room), but knowing that much doesn't clarify the way in which the system prevents these kinds of loops from happening.

I am far from thinking that the critique by Ransom, Fazelpour and Mole is a critical one for the PP account of attention, and the main focus of this paper will not be a response to their paper. However, they bring to light a very interesting aspect of attention in the PP framework that might not have been studied in enough depth: the way the agent controls (or influences) attention, and the way attention is stabilized and destabilized.

I think that a good scenario to look at is Zen meditation (explained in detail in the following section), because instead of focusing on a stream of ever-changing stimuli (such as one of the two superimposed films), the agent aims to keep an open and relaxed attention, which is a thin balance, easily disrupted by distracting stimuli. Such a scenario puts into the spotlight some subtle patterns in the dynamics of attention that can help us understand the role of agential control of attention in a way that is not external to the PP framework.

In what follows I use the phenomenology (and neuroscience) of Zen meditation (zazen) to explore the workings of attention and the system's tendencies toward stability and instability (autovitiation) in a PP framework. PP is a useful framework to explore what happens during zazen, and the phenomenology of zazen, in turn, can illuminate the role of attention in the PP framework.

3. Zazen

> For the practice of Zen, a quiet room is suitable. Eat and drink moderately. Cast aside all involvements, and cease all affairs. Do not think 'good', do not think 'bad'. Do not administer pros and cons. Cease all the movements of the conscious mind, the gauging of all thoughts and views. Have no designs on becoming a Buddha. — Dogen (in Norman and Masao, 2002, p. 3)

Eihei Dogen (1200–1253) was the founder of the Sōtō School of Zen in Japan. His essay 'Fukanzazengi' is probably the most famous instruction on how to sit zazen, a form of meditation that is the central practice of the Sōtō School. I thought that a direct quote of the 'Fukanzazengi' would be more accurate than my own description of zazen, and it would also transmit some of the flavour of Zen writing to the reader.

As I will later refer to my own experience in zazen, I think it is important to relate (very briefly) my zazen practice. In 2014, I visited Kokenji (in the region of Akita in Japan), a Sōtō temple that follows the standards set up by Dogen. For five days I stayed at Kokenji and at its zendo (the meditation hall, high up in the hills), learning from the abbot, Master Sato, about Zen and about zazen. I have been practising zazen every morning (almost) ever since, and last summer I returned to Kokenji for a longer stay, to participate in the practices of a Zen temple, which include zazen, sutra chanting, and daily work (which in Zen is considered part of the spiritual practice).

To have a better understanding of what zazen is, I would like to direct the reader's attention to Dogen's words, 'have no designs on becoming a Buddha' (*ibid.*). These words highlight a big difference between zazen and other methods of meditation: zazen is not an instrumental meditation. It is not the aim of zazen to reach enlightenment, or to become more attentive, or more compassionate. Zazen is not a means to an end; it is an end in and of itself. Although some of the articles referenced in this paper might use expressions such as 'the aim of zazen', or 'the meditator's goal', I believe that it is more productive to think of zazen as an enactment practice.

Having a goal in mind hinders the experience of zazen. This is what Dogen is referring to when he asks, rhetorically, 'how do you think of not-thinking?' (*ibid.*, p. 4). The practice of zazen is to redirect attention away from distractions, including one's own thoughts, and to be fully present in the moment. Thinking about non-thinking still involves thinking.

To outline some of the characteristics of zazen and to clarify where it differs from other meditation practices I will use the taxonomy introduced by Dahl, Lutz and Davidson (2015). Dahl *et al.* review the existing meditation studies to categorize different styles of meditation through the lens of cognitive neuroscience. They divide the different styles into the 'attentional', 'constructive', and 'deconstructive' families. The main characteristic of the 'attentional' family is its focus on the regulation of attention. A prime example of this family of meditation is the practice of focusing on breathing, which is common to many branches of Buddhism. The 'constructive' family aims to strengthen certain mental patterns that improve well-being. The practices of the 'deconstructive' family 'aim to undo maladaptive cognitive patterns by exploring the dynamics of perception, emotion, and cognition and generating insights into one's internal models of the self, others, and the world' (*ibid.*, p. 519).

Dahl *et al.* classify zazen (they use the term 'shikantaza', which I will consider synonymous with 'zazen' for the purposes of this paper) as belonging to the 'deconstructive' family. They sub-classify it as a 'non-dual oriented insight', a type of meditation designed to elicit an experiential shift 'in which the cognitive structures of self/other and subject/object are no longer the dominant mode of experience' (*ibid.*). The taxonomy is based on the primary mechanisms of each practice, but zazen also has some characteristics of the 'attentional' family, as the meditator is supposed to be aware of distractors and not to engage with them. The main difference between zazen and most 'attentional' practices is that in zazen the practitioner doesn't actively direct attention to any given point (e.g. mantras or breathing). Instead, the meditator is just supposed to keep her posture and not to engage with distractors, whether these are external stimuli (a fly buzzing around), bodily stimuli (an itch), or mental stimuli (mind-wandering). When I asked my master how to deal with distracting thoughts he told me, 'Thoughts are always coming and going, do not fight and do not follow. They are like passing clouds in the sky — just let them be' (Joko Sato, in private discussion). I think this is a nice illustration of what it means to sit zazen.

Here is a brief description of one of my early experiences with zazen from my time at Kokenji, which can serve as a representative example of the phenomenology of zazen:

> The sun was setting and the hall was dark. Outside it was raining lightly. From inside, I could hear the drops fall against the roof, and could smell the sweet, wet woods around the Zendo. I found myself thinking about how lucky I was to be in such a place. Maybe I could stay in the Zendo, even. Not go back to Europe where so many small worries had played themselves out before my eyes. I wondered if one day I could become a Zen monk myself.
>
> Thoughts arose answering other thoughts. I suddenly became aware that I was caught in thinking like in a net. This realization only made the thoughts harder to ignore, and as I struggled against them, zazen suddenly seemed insurmountably difficult. My mind didn't like empty spaces. I focused on the tatami in front of me, forcing myself to let go of my thoughts and repeatedly finding myself thinking. Sometimes a dim light on the tatami would catch my eye and I would think how much its movement resembled dancing. Sometimes I would find myself focusing on the sound of the rain. Sometimes sudden excitement came to me; brilliant ideas appeared and I longed to write them down. I suppressed the urge, suppressed the subsequent frustration, and eventually, I was able to let my thoughts go more easily.

Now, after two years of practice, I get 'caught up' in thoughts less often and, mostly, I don't get very frustrated when I do. Although my zazen practice has evolved from my months at Kokenji, the presence of distraction has not abated. As a zazen practitioner, I experience constant distractions that pull my attention away from a relaxed and homogeneous state of attention. In my experience, these are not only my own thoughts, but auditory and visual distractions as well (in the above example, the light on the tatami, or the rain on the roof).

Integral here are the difficulties associated with maintaining a relaxed state of attention. Novice meditators report that most of the time during their zazen practice thoughts are going in all directions, and it is only very experienced meditators that seem to be at ease in maintaining a relaxed state of attention (Kjellgren and Taylor, 2008).

This is very different from the picture of attention that we had seen in the previous section. According to that picture, the attempt to not attend to distractors should (I suggest) act as another self-fulfilling prophecy, reducing the precision of distractors. If that self-fulfilling loop were all there was to keeping attention away from distractions until the agent 'decides' to switch attention, it would be hard to explain why it is not easier to maintain a relaxed attention during zazen.

A popular term in Zen is the 'monkey-mind', the idea that the mind has the tendency to switch its attention from here to there like a monkey swinging in the branches. Although the experience of zazen highlights this tendency of attention to switch objects and to engage with distractors, descriptions of this tendency of attention are not unique to Zen. In his *Treatise on Psychological Optics*, Helmholtz tells the reader that

> the natural unforced state of our attention is to wander around to ever new things, so that when the interest of an object is exhausted, when we cannot perceive anything new, then attention against our will goes to something else. (Helmholtz, 1860, p. 770; translated in Hohwy, 2012)

And precisely referring to the role of attention in PP, Clark describes how hard it is to maintain attention fixed on a single point:

> Try to attend long and hard to a single word on this page. The experience, or so it seems to me, is initially one of increased local clarity, closely followed by a state of decaying clarity while remaining alert. There is at that point a tendency to entrain action, perhaps using shifts of covert attending or micro-saccades to further explore the fixated word. The longer all this goes on *without* the emergence of any new, different, or clearer information the harder it becomes to sustain the process of attending. (Clark, 2015, p. 66)

Is there anything in the PP literature that could explain both this possibility of sustaining attention for short periods of time and the difficulty of sustaining attention for long periods of time? In a paper from 2012 Friston *et al.* study the dynamics of an organism changing from one perceptual state to another. They develop certain assumptions of PP mathematically (namely, prediction error minimization through a hierarchical, non-linear, dynamic generative model) and find that conditional expectations need to maintain uncertainty for a flexible and veridical representation of the world, which leads to their claim that organisms that use PP models will show what they term *autovitiation*: the predisposition of self-organizing systems to destroy their own fixed points (Friston, Breakspear and Deco, 2012).

Intuitively, autovitiation makes sense because the world itself is uncertain and full of noise. Good hypotheses about this world would also include uncertainty to match the uncertainty of the world. Assigning 100% probability to anything is not a good strategy because we only have limited information about the world. It also makes sense that natural selection would benefit agents that don't get 'stuck' in any given hypothesis. There has to be a balance between focusing on something and checking the environment, to make sure, for example, that a tiger is not around when you are picking berries.

There are, of course, mechanisms within the PP framework (and less exotic than autovitiation) that could explain attention switching. One possible explanation is that one of the higher-order levels of the hierarchy causes a change in what the relevant context is, which raises the precision expectations of the input that's relevant to the new context (this was suggested by Clark in private discussion). I think it is useful to think of autovitiation as a property of the system, not as a mechanism. The predisposition of self-organizing systems to destroy their own fixed points might be instantiated through different mechanisms, such as changing beliefs of what the relevant context is for the agent. Autovitiation is a description of the dynamics of a system, for which mechanisms such as stochastic resonance ('by which random fluctuations in a system's state enables it to move over energy barriers and explore multi-stable landscapes'; Hohwy, Roepstorff and Friston, 2008) can serve as an explanation.

Autovitiation is a term from the study of the state-dependent dynamics that underlie PP, but the claim of autovitiation doesn't have a specific commitment to the mechanisms that implement those dynamics. When introducing the relevance of the analysis of dynamic systems to the study of the mind, Kelso is careful to point out that 'dynamic patterns and pattern change are somehow independent of the stuff that realizes them and the level at which they are observed' (Kelso, 1997, p. xiii). The next step here is to bring the discussion back to attention and precision expectations, in a way that is coherent with both the dynamics of self-organized systems and the phenomenology of zazen.

When we look at autovitiation in terms of prediction-making, I believe that it is helpful to think of the system as embodying a hyper-prior predicting a given level of uncertainty. Therefore, when the system is stuck in a given perceptual state that reduces the uncertainty below that level, this generates prediction error, and attention switches to reduce that error by assigning high precision expectation to other perceptual states (distractors). In other words, attention has the tendency to assign higher precision to distracting input (because of the ingrained prediction of uncertainty) in as much as the self-organized system has the tendency to destroy its own fixed points (because of autovitiation). Examples of the equivalent to autovitiation (or the effect of the 'uncertainty hyperprior' at the level of predictions) at the phenomenological level are the difficulty of focusing on the tatami during zazen or Clark's earlier case of trying to attend long and hard to a single word on this page.

Hyperpriors are very firm priors at high levels of abstraction, as in the example we saw earlier with the 'hollow-mask illusion', where the hyperprior that faces are convex impedes the agent from perceiving a concave mask as concave. Statistically speaking, hyperpriors are prior beliefs about parameters (such as a variance parameter) of priors (Friston, Lawson and Frith, 2013). If the 'uncertainty hyperprior' story is right, it is this hyperprior that causes the system to assign higher precision to distracting input (through the mechanism of attention as defined by Hohwy). Distracting input, in PP terms, would be input that challenges the current state of conscious perception (which, in Hohwy's definition, is determined by the prediction with the highest overall posterior probability). If the current state of conscious perception were stabilized for too long, the system would assign lower precision to the input confirming the current perceptive state and higher precision to the error challenging the current perceptive state, because this error would be consistent with the hyperprior that the world has a certain level of uncertainty and it would thus be considered more reliable input.

4. Cognitive Control

The mind is constantly being assailed with distractions during zazen, and yet it is possible not to engage with those distractions, to let them go. It is not, of course, always possible, and the extent to which it is possible depends on many factors, one of the most important ones being the number of years of meditation experience. What are the mechanisms that enable the meditator to cease 'the gauging of all thoughts and views' (Dogen in Norman and Masao, 2002, p. 3) and how — if at all — do they fit in the picture of attention that we have been exploring?

The ability to control our thoughts (and our attention) in line with our intentions is known as cognitive control, and it 'is closely tied to the prefrontal cortex (PFC), which is proposed to harbour temporary representations of current goals, goal-relevant stimuli and strategies' (Jiang, Heller and Egner, p .31). As one could expect, studies using fMRI show that the PFC 'is significantly affected by meditation, with experienced meditators showing improved efficiency and response inhibition relating to these areas' (Larson, Steffen and Primosch, 2013, p. 2).

There are some problems making generalizations from studies on other meditation styles to zazen, because very few neural and behavioural studies on meditation use the fine taxonomy introduced earlier in this paper (Dahl, Lutz and Davidson, 2015). Most studies make a coarse distinction between two fundamental forms of meditation: focused awareness (FA) and open monitoring (OM). During FA the subject's attention is centred on a particular object (e.g. breathing), while OM aims for a non-reactive awareness of the contents of consciousness in the present moment. Although most meditation styles share characteristics of both FA and OM, zazen is clearly closer to OM than to FA. Non-reactive awareness echoes Dogen's expression 'Do not think "good", do not think "bad". Do not administer pros and cons' ('Fukanzazengi'). In what follows, I will look at evidence coming from studies not only on zazen, but also on other types of meditation (primarily OM). Of course, any extrapolations from non-zazen OM meditation styles to zazen should be done carefully, and tested empirically whenever possible.

While studies of FA show activity in areas related to attentional focus, OM is associated with 'brain regions involved in vigilance, monitoring and disengagement of attention from sources of distraction from the on-going stream of experience' (Manna *et al.*, 2010, p. 47). This disengagement from distraction matches our previous analysis of the phenomenology of zazen, and it is not only present in phenomenology and neuroscience, but in behavioural tests as well. In a test for sustained attention 'OM meditators showed superior performance… in comparison with FA meditators when the stimulus was unexpected; however, there was no difference between the two groups of meditators when the stimulus was expected, indicating a more distributed attentional focus in the OM meditators' (experiment from Valentine and Sweet, 1999; quote from Lutz *et al.*, 2008, p. 165).

At the level of the dynamics underlying these mental processes, the claim (corresponding to what we have seen at a neural, behavioural, and phenomenological level) would be that the meditators' intentions affect autovitiation in such a way that the incoming attractors (in the form of distractions) get destabilized. Scholz and Kelso (1990) devised an experiment in which the subject's intentions affected the dynamic patterns of their actions. The subject's initial task was to cycle their fingers in phase or antiphase. An auditory signal would indicate to them to switch to the opposite coordination mode as fast as they could. Their prediction was that faster switching should happen from less stable to more stable patterns (i.e. follow the subject's intrinsic dynamics) and that intentions could change the dynamical stability of the patterns. The experimental results matched the calculated theoretical distribution of switching times for both directions of switching given these predictions. Kelso concludes that intention (in this case, the subject's intention to change the coordination mode in which she cycles her fingers) 'acts to parameterize the intrinsic dynamics, stabilizing one pattern and destabilizing the other… Yet the presence of the intrinsic pattern dynamics… is always felt'. Thus, 'intentionally sustaining intrinsically unstable patterns has costs' (Kelso, 1997, p. 152). The effects of intentions on action help us illustrate the hypothesized effects of intention on attention during zazen. If the above picture is correct, the intention of not engaging with distractors destabilizes the incoming attractors (i.e. distractions), and stabilizes the current state; a process that is costly and challenging for the meditator.

One big difference between attention during zazen and the Scholz-Kelso experiment is that zazen doesn't involve a one-time intention, but a longer-term 'strategy' (i.e. not to engage with distractors during zazen). This strategy (of the kind hypothesized to be maintained in the PFC) would have an effect not only on the individual decisions not to engage with distractors, but in the overall dynamics during meditation. I believe that, in terms of prediction-making, the zazen meditator uses cognitive control to try to maintain a globally distributed precision expectation to cancel out the influence of the 'uncertainty hyperprior' on attention. This globally distributed precision expectation corresponds to the non-reactive awareness of the contents of consciousness showed in most OM meditation practices, and the relaxed state of attention (i.e. globally distributed attention) of experienced Zen practitioners.

There is an interesting 'decoupling' of the self during zazen. There seems to be a self who wants to engage with distractors (following a hyperprior of instability; felt as a pull towards the distractor during zazen experience), and the self who wants to disengage from distractors (following a global prediction of stability; felt as a smoothing out of the pull of distractors). The way in which the agent's intentions affect attention is by taking the decision of sitting zazen (a global strategy of disengaging with distractors). However, if the meditator ends up engaging with a given distractor (e.g. following a stream of thought), it could be experienced as her decision to engage with that distractor.

In an experiment on volitional acts, Libet found that 'freely voluntary acts are preceded by a specific electrical change in the brain (the "readiness potential", RP) that begins 550 ms before the act. Human subjects became aware of intention to act 350–400 ms after RP starts, but 200 ms before the motor act. The volitional process is therefore initiated unconsciously. But the conscious function could still control the outcome; it can veto the act' (Libet, 1999, p. 47). RPs have also been recorded before attention switching (Luck, Woodman and Vogel, 2000). It seems likely that, during zazen, a pull to engage with a distractor triggers an RP, and that this RP can then be controlled or vetoed through cognitive control.

The Libet experiment is useful to illustrate that, whether the meditator engages or disengages with the distractor, she could experience it as a decision. In the Libet experiment, the subjects described the decision to act (e.g. to press a button) when they acted, and the decision not to act (e.g. not to press the button) when they vetoed an impulse to act (preceded by an RP). There is, of course, a complex interplay of factors that leads the meditator to engage or disengage with a distractor. These include the influence of the instability hyperprior versus the globally distributed precision expectation, but not only this. There are many priors involved at any given moment, and some distractors pull attention more than others (dependent, for example, on the level of noise of the incoming errors). The capacity of cognitive control is different in different individuals, and this also plays a role in the way that attention is distributed.

We have, out of zazen, a very coherent picture of how attention works, which can be extrapolated to realms outside of zazen meditation: attention is the process of precision optimization, which biases conscious perception by giving more weight to the input that has higher estimated precision. This estimation of precision depends on the complex interplay of many different factors. One of those factors is a hyperprior predicting a given level of uncertainty, which pushes attention to switch targets. Finally, the agent can exert cognitive control by creating a global strategy that can in turn influence the precision expectations.

In return, this clarification of the role of attention makes predictive processing a great framework to conceptualize zazen practice: through cognitive control, the zazen practitioner aims to maintain a globally distributed precision expectation to counteract the distracting effect of a hyperprior predicting uncertainty. This strategy (hypothesized to be harboured in the PFC) is costly for the individual, and it is harder to maintain the more it goes against the intrinsic tendencies of the system. At the personal, phenomenological, level the meditator may experience switching attention as her own decision (or she may not, as with a case in which a surprising input 'grabs' our attention), or non-switching as her own decision, but, at the level of predictions, it all depends on this complex interplay of many different factors. Furthermore, it is possible, through the practice of zazen, to train our cognitive control to influence the workings of attention. An example of this is the way in which expert Zen practitioners are able to maintain a relaxed state of attention.

5. Challenges

One clear challenge to using zazen in cognitive science is that it is an introspective method. Lutz and Thompson identify three main issues with incorporating first-person data into cognitive neuroscience: '(i) first-person reports can be biased or inaccurate; (ii) the process of generating first-person reports about an experience can modify that experience; and (iii) there is an "explanatory gap" in our understanding of how to relate first-person, phenomenological data to third-person, biobehavioural data' (Lutz and Thompson, 2003, p. 33).

Lutz and Thompson look at these challenges from the standpoint of 'neurophenomenology', a research programme that aims to use first-person data 'to describe and quantify the large-scale neurodynamics of consciousness' (*ibid.*). Although I do not share all of the commitments of 'neurophenomenology', their general approach is a useful one for the purposes of this paper. Their general approach, at a methodological level is '(i) to obtain richer first-person data through disciplined phenomenological explorations of experience, and (ii) to use these original first-person data to uncover new third-person data about the physiological processes crucial for consciousness' (*ibid.*, p. 34).

I believe that zazen is a great example of a disciplined phenomenological exploration of experience. Zazen is a well-defined method of exploring the dynamics of our own experience, and one that has been practised and perfected for centuries. One of the advantages of using zazen instead of other methods of phenomenology (such as *epoché*; Depraz, 1999) is that there are a lot of zazen practitioners around the world, and throughout history. Thus, although first-person reports can be biased or inaccurate, it is possible to find patterns in the phenomenology of zazen practitioners and categorize certain aspects of it, which is made easier by the large quantity of data and practitioners available.

The idea of using Buddhist meditative traditions to reveal important phenomena of human consciousness is not new. It is central, for example, to the seminal work of Varela, Thompson and Rosch on the embodied mind (Varela, Thompson and Rosch, 1991). It might have been possible to use a different meditative practice than zazen for phenomenological analysis, and in this paper I have looked at evidence coming from studies not only on zazen, but also on other types of meditation. The reason why I have used zazen is threefold. First, it is the practice that I have been trained in, and in which I have easy access to other practitioners (in particular experienced practitioners to consult with, such as Joko Sato). The phenomenology of my own zazen practice has often guided the different arguments of this paper, and I don't have this kind of phenomenological access to other disciplines. Second, zazen is a very straightforward and unadorned practice. It would be more complex, for example, to use the phenomenology of more complicated meditative exercises such as reciting mantras, or generating compassion, or to disentangle the cultural ramifications of the image-rich meditations of Vajrayāna Buddhism. Finally, zazen is a deconstructive practice (following the taxonomy by Dahl, Lutz and Davidson, 2015); it goes against certain tendencies of the mind, which brings not only those tendencies into focus, but also the difficulty for the subject to go against those tendencies. This is of critical importance for the arguments about autovitiation (i.e. a tendency of the mind) and about the difficulty of sustaining a globally distributed attention (i.e. the subject going against that tendency).

One caveat is that not all zazen practitioners have the same experience. Some practitioners experience mind-wandering more often than others. Some find zazen relaxing and some find it stressful. Even when we look at a single practitioner, her experience of zazen might change depending on the occasion. Although this is something that should be taken into account, I don't consider it a critical problem. As I said, certain patterns are recognizable in the experience of zazen practitioners. There are also known factors that influence the experience of the practitioners. For example, the frequency with which a practitioner sits zazen (how many times per week) affects the experience, and more experienced meditators are not as likely to mind-wander. All of these parameters can be taken into account when comparing the reports of different practitioners.

Finally, research has shown that meditation experience predicts introspective accuracy (Fox *et al.*, 2012), which helps with the concerns about the inaccuracy of first-person reports. Of course, this doesn't cast out the worry of the exercise of zazen modifying normal, everyday experience. Experienced meditators report that they can maintain the type of awareness characteristic of zazen throughout their daily activities. However, in this paper I have not been looking at the phenomenology reported by zazen practitioners in their everyday life (i.e. while they are not sitting zazen), but at the phenomenology *during* zazen, and at what it can tell us about the mind (specifically about attention). Sitting zazen modifies the practitioner's experience, but it is precisely that modified experience (which is zazen) that is of interest for the purpose of this paper.

Regarding the explanatory gap between first-person reports and third-person data, this paper made no attempt to bridge that gap. It was not my ambition to give the same standing to phenomenological reports and to third-person evidence. In the previous section, I have looked at, on the one hand, how phenomenological reports fit in an already existing paradigm, and, on the other, how they can help illuminate issues with the paradigm, and suggested certain revisions. This revision of the workings of attention highlights previously unseen links between different already existing research findings, but they have to, ultimately, be put to test through empirical methods.

Different experiments could explore different aspects of the proposed role of attention, and there might not be a single experiment that can support all of the aspects at once. One such experiment would be to test experienced and non-experienced meditators during zazen, and compare the activity in the PFC with the emergence of ERP, and the reported moments of engagement with distractors. Another possible experiment is to have a case of binocular rivalry (where each of the subject's eyes is exposed to a different image) and measure the intrinsic rates of switching between the two images. Afterwards, the subjects would be asked to maintain their attention on one image, and then 'decide' to switch whenever they want. On a third round, the subjects would switch attention whenever they are told to. I predict two things: 1) when the subjects 'decide' to switch, their switching dynamics would approximate their intrinsic tendency; 2) the further the suggested switching in round three is from the intrinsic dynamics of the subject, the harder it would be for them to switch. A possible version of the experiment would be to then train the subjects in zazen and repeat the test. I predict that they will have become more able to switch images against their intrinsic dynamics.

These experiments wouldn't directly support PP, but if PP were true (I think it is), the experiment sketched would further support my claims about attention. Potentially, several parameters (such as noise) could be added to the experiment to see if they affect the results in ways that are correctly simulated by a Bayesian/predictive model.

6. Conclusion and Future Directions

In the initial sections of this paper I looked at Hohwy's theory of the role of attention within the PP framework, and at some recent challenges for this theory. Looking at the phenomenology of zazen, I proposed that the system embodies a hyperprior predicting a given level of uncertainty. I then looked at the role of cognitive control and how it influences attention, and concluded that attention depends on a complex interplay of different factors, one of which is a hyperprior of uncertainty, and another a global strategy (harboured in the PFC) taken by the agent (e.g. a globally distributed precision expectation), which can influence the working of attention through cognitive control (that can be trained through, for example, zazen). I have then conceptualized zazen within the PP framework. In the last section I have sketched some possible challenges for this conclusion, the main one being the use of introspection.

A matter left to explore is how autovitiation (as a process in system dynamics) *precisely* relates to different mechanisms linked with attention. A possibility is that there is a mechanism involving an endogenous mode of destabilization. For example, a given mechanism (such as stochastic resonance) could periodically create random error signals to destroy the current attractors of the energy landscape and keep the system moving.

Another future direction is to study how attention works for experienced zazen meditators. Tests on brain activity of experienced OM meditators suggest that open monitoring is also reflected and thus practised in ordinary non-meditative conditions (Brefczynski-Lewis *et al.*, 2007). If the mental state of zazen becomes a normal mental state for experienced practitioners, or a habit system, this seems to indicate that their generative models have changed through practice. One possibility is that the expectation of uncertainty leading to distraction has been weakened, but I don't think this the case. Zen Buddhism tries to come to terms with uncertainty (close to the concept of impermanence, which is one of the three Buddhist marks of existence), not to ignore it. It is more likely that the strategy predicting a globally distributed precision expectation has become deeply rooted. Another possibility is that they have developed very efficient cognitive control, so that for them imposing a global strategy is less costly than for inexperienced meditators. These questions should again be put to empirical test, and it might turn out that experienced meditators have both more cognitive control and different generative models.

An outstanding question is how attention is related to other elements of the mind (such as the self), and how this fits in the PP paradigm. Finally, one could wonder what the purpose of zazen is in reinforcing cognitive control to counteract the mind's tendencies, especially if such tendencies are useful for correctly predicting the world around us, and are beneficial for natural selection. With that question in mind, I conclude with the following quote from Dogen:

> We study the self to forget the self. And when you forget the self, you become one with all things. (Dogen, translation in Snyder and McLean, 1980, p. 65)

Acknowledgments

Thanks to Andy Clark for his helpful commentaries, his guidance, and his sharp, precise, and constructive criticism. Thanks to Joko Sato, for his invitation to stay at Kokenji Zendo, his patience, and his unswerving example. Thanks to María Velasco for her faith and her support.

References

Brefczynski-Lewis, J., *et al*. (2007) Neural correlates of attentional expertise in long-term meditation practitioners, *Proceedings of the National Academy of Sciences*, **104**, pp. 11483–11488.

Clark, A. (2015) *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*, Oxford: Oxford University Press.

Dahl, C.J., Lutz, A. and Davidson, R. (2015) Reconstructing and deconstructing the self: Cognitive mechanisms in meditation practice, *Trends in Cognitive Sciences*, **19** (9), pp. 515–523.

Depraz, N. (1999) The phenomenological reduction as praxis, *The View from Within: First-Person Approaches to the Study of Consciousness*, special issue of *Journal of Consciousness Studies*, **6** (2–3), pp. 95–110.

Fox, K.C., Zakarauskas, P., Dixon, M., Ellamil, M., Thompson, E. & Christoff, K. (2012) Meditation experience predicts introspective accuracy, *PloS one*, **7** (9).

Friston, K. (2010) The free-energy principle: A unified brain theory?, *Nature Reviews Neuroscience*, **11** (2), pp. 127–138.

Friston, K., Breakspear, M. & Deco, G. (2012) Perception and self-organized instability, *Frontiers in Computational Neuroscience*, **6**, art. 44.

Friston, K., Lawson, R. & Frith, C. (2013) On hyperpriors and hypopriors: Comment on Pellicano and Burr, *Trends in Cognitive Sciences*, **17** (1), art. 1.

Friston, K., *et al*. (2016) Active inference: A process theory, *Neural Computation*, **29** (1), pp. 1–49.

Hohwy, J. (2012) Attention and conscious perception in the hypothesis testing brain, *Frontiers in Psychology*, **3**, art. 96.

Hohwy, J., Roepstorff, A. & Friston, K. (2008) Predictive coding explains binocular rivalry: An epistemological review, *Cognition*, **108** (3), pp. 687–701.

Jiang, J., Heller, K. & Egner, T. (2014) Bayesian modeling of flexible cognitive control, *Neuroscience & Biobehavioral Reviews*, **46**, pp. 30–43.

Kelso, J.S. (1997) *Dynamic Patterns: The Self-Organization of Brain and Behaviour*, Cambridge, MA: MIT Press.

Kjellgren, A. & Taylor, S. (2008) Mapping zazen meditation as a developmental process: Exploring the experiences of experienced and inexperienced meditators, *Journal of Transpersonal Psychology*, **2**, pp. 224–250.

Larson, M.J., Steffen, P.R. & Primosch, M. (2013) The impact of a brief mindfulness meditation intervention on cognitive control and error-related performance monitoring, *Frontiers in Hhuman Neuroscience*, **7**.

Libet, B. (1999) Do we have free will?, *Journal of Consciousness Studies*, **6** (8–9), pp. 47–57.

Luck, S.J., Woodman, G.F. & Vogel, E.K. (2000) Event-related potential studies of attention, *Trends in Cognitive Sciences*, **4** (11), pp. 432–440.

Lutz, A. & Thompson, E. (2003) Neurophenomenology: Integrating subjective experience and brain dynamics in the neuroscience of consciousness, *Journal of Consciousness Studies*, **10** (9–10), pp. 31–52.

Lutz, A., Slagter, H., Dunne, J. & Davidson, R. (2008) Attention regulation and monitoring in meditation, *Trends in Cognitive Sciences*, **12** (4), pp. 163–169.

Manna, A., Raffone, A., Perrucci, M., Nardo, D., Ferretti, A., Tartaro, A. & Romani, G.L. (2010) Neural correlates of focused attention and cognitive monitoring in meditation, *Brain Research Bulletin*, **82** (1), pp. 46–56.

Neisser, U. & Becklen, R. (1975) Selective looking: Attending to visually specified events, *Cognitive Psychology*, **7** (4), pp. 480–494.
Norman, W. & Masao, A. (2002) *The Heart of Dôgen's Shôbôgenzô*, New York: Suny Press.
Ransom, M., Fazelpour, S. & Mole, C. (2016) Attention in the predictive mind, *Consciousness and Cognition*, **47**, pp. 99–112.
Scholz, J.P. & Kelso, J.A.S. (1990) Intentional switching between patterns of bimanual coordination depends on the intrinsic dynamics of the patterns, *Journal of Motor Behavior*, **22** (1), pp. 98–124.
Snyder, G. & McLean, W. (1980) *The Real Work: Interviews & Talks*, New York: New Directions Publishing.
Valentine, E.R. & Sweet, P.L. (1999) Meditation and attention: A comparison of the effects of concentrative and mindfulness meditation on sustained attention, *Mental Health, Religion & Culture*, **2** (1), pp. 59–70.
Varela, F., Thompson, E. & Rosch, E. (1991) *The Embodied Mind: Cognitive Science and Human Experience*, Cambridge, MA: MIT Press.
Wilkinson, S. (2014) Accounting for the phenomenology and varieties of auditory verbal hallucination within a predictive processing framework, *Consciousness and Cognition*, **30**, pp. 142–155.

Paper received March 2017; revised September 2017.