

The Foreground and Background of Consciousness – An Introspective Argument Against Introspection

As for the pure philosophical 'freedom of the will' my will is as free as I feel it to be and there is an end to the matter.

– R.B. Braithwaite.¹

There is a widespread tendency in the philosophy of our times to 'get back to basics'. In the wake of philosophical reconstruction workers such as G.E. Moore and Ludwig Wittgenstein, and the ordinary language philosophy they inspired, contemporary philosophers appear to have learned some important lessons: to take into account our most commonsensical notions about ourselves, each other, and the world we live in, and to beware of framing elaborate problems in theory where practically there are none. The result of such a common sense mentality is not only that our ordinary phenomenological apprehensions of life in general are taken seriously – as they should be, if philosophy is to make sense of the world we live in – but also that certain assumptions about our self-experience are allowed to go unquestioned. Thus we see that in many philosophical texts, especially those discussing free will, consciousness, agency, and deliberative awareness, a specific phenomenological² account of the self is presupposed: namely, the self as a fixed entity, a unified centre of consciousness, which focuses its intentions into actions, and hence experiences itself as the source of its actions. Such conceptions are usually based on philosophers' armchair introspection, yet they are deemed to be universal among mankind, an undeniable part of our most basic experience and common sense. Indeed, what can be more commonsensical than that we experience ourselves as selves?

I intend to question this assumption, particularly in the context of the free will debate. I do believe that such a perception of the self may arise when we put on our 'introspectacles': that is, when we are consciously thinking or talking about ourselves, and when we philosophically focus on 'the self itself'. However, I do not believe that it is warranted on the mere basis of such introspection to infer that this is *how we usually experience ourselves*. Could it not be the case that the often-reported experience of the 'self' and some of its attributes is a *result* of focused introspection, rather than an expression of our default self-experience? Perhaps the experienced 'self' *prior* to introspection encompasses a wider and more dynamic background of conscious, semi-conscious and preconscious aspects, while introspection focuses on what is at the foreground of our minds, and translates this into the experience of a consciously willing self.

This paper will focus on the nature of the introspective move itself, on the interpretation that may be implicit in such a moment of 'looking into oneself'. The main question to be answered is as follows: is the 'self' we encounter in introspection phenomenologically identical to how we *usually* experience ourselves? In what follows I will first of all survey the predominance of this 'introspective self' in the free will debate, and briefly discuss some previous criticisms of philosophers' use of phenomenology, in order to prepare a deeper critique of introspection. Second, I will try to expose several problems stemming from the nature of introspection itself, and argue that it is at least *possible* for the 'introspective self' to differ from the default 'experiential self'. Third, I will – perhaps paradoxically – call on introspection for evidence that such a divergence between introspected and experienced selves is phenomenologically plausible, since a different kind of introspection may lead to different introspective results. To eluci-

date this ‘introspective argument against introspection’ I will introduce the concepts of a foreground and background of the experience of consciousness. Finally, I will conclude that it is time to further problematise the use of introspection in philosophy, and briefly consider the implications of these arguments for the free will debate.

1. The ‘Phenomenal Fact’

While there is no scarcity of phenomenological universals³ concerning the nature of the self in any type of philosophy, it is in the context of the free will debate that they make their most prominent appearance. The free will problem can be construed as ‘an unexplained gap between the category of physical phenomena and the category of subjective phenomena’ (Libet, 1999: 55) – that is, as a discrepancy between the implausibility of mental causation (i.e. that our mental states can cause physical events – see Hohwy below) from a metaphysical point of view, on the one hand, and our experience of ourselves as conscious agents, on the other.⁴ To put it more bluntly: philosophy may tell us we are not free, but ‘we’ feel we are. Solutions to the problem usually consist of either an attempt to bridge the metaphysical-experiential gap, or an effort to show that one of the two conflicting terms should be prioritised. Either way, the notion of a unitary, consciously acting self is invoked by many proponents of mental causation in order to voice an innermost experience, to which we all are supposed to be committed. At the same time, this notion is taken seriously by opponents of mental causation, who likewise consider it a basic part of our experience, be it one that must be explained away. Both sides of the debate, then, broadly seem to agree that the introspective experience of the acting self is a hard ‘phenomenal fact’ (to borrow a phrase from Libet, below), yet they differ in the value they assign to it: whether as a piece of ‘*prima facie* evidence’, or as an illusion that should not be given philosophical credence.⁵

Consider the appearance of this ‘phenomenal fact’ in a variety of papers (I have highlighted the universalising tendencies in bold):

‘[...] we must recognize that **the almost universal experience** that we can act with a free, independent choice provides a kind of *prima facie* evidence that conscious mental processes can causatively control some brain processes [...] The phenomenal fact is that **most of us feel** that we do have free will [...]’ (Libet, 1999: 56).

‘[...] the agency theory is appealing because it captures **the way we experience** our own activity. It does not seem to me (at least ordinarily) that I am caused to act by the reasons which favour doing so; it seems to be the case, rather, that I produce my own decisions in view of those reasons [...]’ (O’Conner, 1995: 196).

‘[...] **your phenomenology** presents your own behavior to you as having yourself as its source, rather than (say) presenting your own behaviour to you as having your own occurrent mental events as its source [...]’ (Horgan et al., quoted in Nahmias *et al.*, 2004: 167).

‘**Many people, including most philosophers**, have a very firm belief that there is mental causation, that is, that mental states such as beliefs, desires, intentions, and emotions are efficacious [...] in the causing of some physical events such as bodily movements and actions in the wider sense [...] **We thus have a very deep attachment** to mental causation’ (Hohwy, 2004: 377).

‘[...] it seems **to each of us** that **we** have conscious will. It seems **we** have selves. It seems **we** have minds. It seems we are agents. It seems **we** cause what **we** do’ (Wegner, 2002: 342).

‘Human freedom is just **a fact of experience**. [...] a series of powerful arguments based on facts of **our own experience** inclines **us** to the conclusion that there must be some freedom of the will because **we all experience it** all the time’ (Searle, 1984: 88).

What emerges from an overview of such ‘low-brow’ accounts of phenomenology – and many other examples can be found – is an informal

yet introspectively plausible sketch of the thinking, willing, acting self. The ‘actish phenomenal quality’ (Ginet, 1990: 13) at the source of these introspective musings springs from individual philosophers, yet their conclusions are universalised in a subtle move from ‘I’ to ‘we’, ‘each of us’ or ‘most of us’.

Such practices of introspection or first-person phenomenology have been subjected to a variety of criticisms in the past decades, which in turn have given rise to new methods and versions of phenomenological investigations. For instance, Eddy Nahmias *et al.* have criticised the ‘universality assumption’ of free will philosophers who believe their own (often mutually incompatible) introspective reports are indicative of the human condition in general (2004: 164). Instead, the authors propose ‘systematic psychological research on the relevant experiences of non-philosophers’ (169) – that is, they attempt to gain insight to lay phenomenology through a series of empirical queries, and conclude that the ‘universality assumption’ is often wrong. Philosophers usually ‘introspect through the lens of their theoretical commitments’ and are therefore ‘the wrong subjects to trust’ (163).

Daniel Dennett appears to be expressing a similar concern in his critique of ‘the *first-person-plural presumption*’: when thinking about consciousness, ‘I’ is often widened into ‘we’ (1991: 67). Not only is this presumption unwarranted, says Dennett, but our very notion of infallible introspection is a mistake: ‘I suspect that when we claim to be just using our powers of inner *observation*, we are always actually engaging in a sort of impromptu *theorizing* [...]’ (*Ibid.*). There is no way of eliminating interpretation from such observation – thus all (not just philosophical) introspection is suspect. Yet Dennett does believe it possible to construct an objective method of phenomenology, which avoids the temptations of first-person introspection without sliding into a fully reductive behaviourism that forbids any talk of mental events. This ‘heterophenomenology’ describes subjects’ reports in the scientific third-person perspective, and treats the reported intentional objects like fictional entities, which may or may not be real (*Ibid.*: 71-98).

These alternative methods of phenomenology, and others of the sort, share common ground in that they question the salience of individual

introspection and tend towards a broader scientific survey of human experience. Herein lies an important insight. Yet these methods do nothing to answer my question about what happens in the introspective move: they criticise phenomenological claims just with respect to their *universalisation* of introspection, not with respect to introspection itself. The more fundamental question about introspection, I would argue, cannot be evaded, even by Nahmias and Dennett. If we want to avoid radical behaviourism and be able to say anything about consciousness or experience, as both authors do, somewhere along the line *someone* has to introspect – whether it be the philosopher or the subjects of an experiment. The Nahmias queries, though distanced from philosophers’ introspection, remain dependent on the lay subjects’ introspection. Dennett’s heterophenomenologist is merely re-describing the first-person experiences of the subject into a third-person theory. (The scientist who notes that ‘subject S reports having an experience E’ may not have to introspect personally, yet still relies on S introspecting and commenting on E.) Either way, empirical queries of phenomenology cannot entirely steer clear of the first-person perspective. Hence, the ‘deeper’ question concerning introspection has not yet been answered. On the contrary – it has not even been posed.

2. The Phenomenological Fallacy

The deeper problem, remember, is whether introspection adequately represents our common experience of ourselves to ourselves. To argue that such an assumption of adequacy is at least mildly suspicious and perhaps even highly problematic, I will make two remarks – one might call them premises – on which to build my case. Since my argument is designed to make a general point, it should not depend on a specific theory of either phenomenology or introspection, and I do not want to commit to one. I will therefore not introduce detailed theoretical and methodological underpinnings of the topics in question, and merely outline some points that should be *relatively* modest and uncontroversial.

The first point should be clear from the discussion above: phenomenology is not like most ‘normal’ sciences. It has for its subject matter not the outer world as it is, but as it appears to the inner world. As the case of

Dennett shows, even if we aspire to objectively describe the phenomenon, in phenomenology we are ultimately bound to some kind of subjective experience: to the first-person perspective. If phenomenology wants to investigate consciousness and experience – what it is like to be or see or feel something – it has to rely on the ‘inner take’ on things. It cannot naively fall back on the scientific ‘outer take’ without forfeiting the project itself, as is the common tendency in behaviourism. Phenomenology without a *view* from the inside, I would argue, is not phenomenology at all.

The basic subjective commitment of phenomenology opens a realm of possibilities as well as problems of its own. For even if it is agreed that we cannot access subjective experience from the outside without losing an essential part of this experience (the ‘what it is like’-ness) – how do we know we can do justice to the experience while describing it *from the inside out*?

This is where introspection comes in, and it brings us to the second point: introspection is not like ‘normal’ perception or observation. The subject-object vocabulary that makes sense when we are discussing our perception of things becomes problematic in introspective contexts, where the phenomenon seems to be observing itself. According to a ‘direct access’ (or ‘unmediated observation’) model of introspection, this lack of differential distance between subject and object is precisely what endows introspection with an immediate and infallible source of knowledge.⁶ Following the footsteps of Descartes (and perhaps Husserl), one might posit the absolute transparency of the self, and infer from this that the most certain knowledge can be derived from introspection.

I believe this model of introspection is the only one that is fundamentally incompatible with my argument below. Since it is widely discredited, I will permit myself to dismiss it on two grounds. First, there appear to be no solid arguments for the notion of introspective self-transparency except perhaps the ‘*prima facie* evidence’ of introspection itself – which is exactly where the problem lies. Second, starting from the premise that there is no subject-object differentiation in introspection as there is in perception, one could argue in either of two ways. One might follow the direct access model in arguing that, wherever subject and object are the same, there is an immediate and therefore certain route to knowledge. But one could

just as well take the opposite direction, and argue that where there is no separate object, there is no objectivity, and so the absence of distance is precisely what makes introspection problematic where perception is not.⁷ It is this second line of reasoning that I would like to pursue a bit further.

It is possible to argue for any kind of observation that the act of looking changes the object observed. This worry could arise at different levels: in the Kantian consideration that we can never experience the *Ding an sich*, the thing-in-itself that is out of reach of the senses; or in the more trivial notion that we only ever see the world through our own eyes. It rises in a different way in quantum physics: for instance, in Schrödinger’s ‘cat in the box’-scenario, where it is the very act of looking that determines whether the cat is dead or alive. We could argue, along the transparency-line, that the absence of a basic subject-object distinction makes introspection immune from worries of this kind. Yet we could also argue that the worry becomes more fundamental in this context, as the act of introspection may set up a new and more problematic subject-object distance *within the self*: i.e. the observing-self (subject) versus the self-observed (object). The classic dilemma is whether introspection works through a kind of immediate transparency, or whether it creates a self-as-object that is somehow different from the self-as-subject of our experience.

I do not believe this problem can be solved *a priori* – if it can be solved at all. Fortunately, I do not need to solve it to be able to draw the following conclusion: it is at the very least conceptually and metaphysically *possible* that introspection is not immediate, but is itself a kind of mediation. This possibility is all that is required to present some of the free will philosophers quoted above with the following objection: that they implicitly *and unwarrantedly* assume that their phenomenology and introspection are somehow immediate and therefore constitute a kind of ‘*prima facie* evidence’ to support conclusions about a variety of mental phenomena. I call this the ‘phenomenological fallacy’: it consists mainly of speaking too easily of ‘phenomenal facts’. Considering the possibility of introspective mediation, such implicit trust in active introspection may be dangerous as well as unwarranted. Again, the question must be asked whether the introspected self does justice to the full range of our self-experience.

Here the reader might tug my sleeve, as I appear to be sliding into

dangerous territory. Surely I am not positing an experienced self versus a *real* self – or, to speak in Kantian terms, a phenomenal versus a noumenal self (the self ‘*an sich*’)? On the contrary. I would like to consider the possibility of two levels of ‘phenomenal’ selves: the self as we encounter it in introspection proper, as opposed to the way we usually experience ourselves. I am interested not in how the self *really* is – after all, there is not much we can say about that – but in how it *usually* is experienced, when we are not actively, philosophically, introspecting.⁸ That it is possible for the introspective and experiential self to come apart may not seem intuitively plausible at first. Yet I believe it is possible to realise *through introspection itself* how introspection may distort our self-experience. This may *seem* paradoxical, but it is also unavoidable. That is: we cannot proceed much further here by way of mere argument. Beyond the possibility of mediation mentioned above, we have come to the limits of what we can say about the nature of introspection without invoking some kind of introspection itself. In other words, it is time either to be silent or to introspect. I believe it is worth exploring the second option.

3. The Introspective Argument Against Introspection

Let’s take a step backwards. Perhaps we can reconstruct the introspective technique on which the ‘common sense’ phenomenology of the self and the will is based – the kind that I suspect lies at the heart of the phenomenological fallacy – as follows:

[INTROSPECTION 1]: Look into yourself. What do you see? You see and feel yourself as an entity of a certain character, with a specific set of desires and beliefs, an irreducible ‘I’ that is separate from the world. Based on what you want to do and what you think is best to do, you make certain decisions, and realise your intentions into actions. Doing so, you have a ‘feeling of doing’: you are aware that you have consciously willed these actions, and hence you experience yourself as an agent. Look back at your past actions: usually, when you did something, you intended to do it and then did it. Look forward to your future actions: nothing

is stronger than the feeling that it will be *you* who is acting, not merely your body or your brain. Raise your right hand. Did you not feel it was *you* who raised it?

Is this an exaggeration? If it is, it is not meant to be. In the context of this paper, one may already be inclined to take a critical stance to such a leading introspective exercise. But introduced and worded in the right way, such invocations of ‘low brow’ introspection can have a strong intuitive appeal. To me personally, at least at first sight, the results of [INTROSPECTION 1] seem quite convincing. In the light of actions performed, it seems only natural to think of my agency along the lines of willing-doing: I wanted to go for a walk in the evening, so I decided to do it, and then I did it. Thinking of the future, I can already frame intentions in my mind that will lead to the actions intended. Most strikingly, any version of the hand-example can have the common sense plausibility of Moore’s paradigmatic ‘Here is one hand’- approach.⁹ When I introspect while observing my hand and deciding whether or not to raise it, and then raising it, there does seem to be a ‘free-willish’ feeling of doing accompanying my movement. Nothing seems more plausible than to say that this is a hand, and this is me raising it.

In themselves, I believe auto-experiments of this kind, in their appeal to common sense experience, are based on a good hunch, a healthy philosophical instinct. But they fail here, since common sense is precisely what is at stake: what *is* our everyday, pre-theoretical experience of ourselves and our actions? What if the kind of self-reflection invoked in [INTROSPECTION 1] is already an interpretation? Here we would do well to remember Wittgenstein’s criticism of Moore’s commonsensical examples: when philosophers say things like ‘I am certain here is a hand’ or ‘that is a tree’, they are already far removed from ordinary language and everyday cases (Wittgenstein, 1969/1975). Similarly, there is no ordinary context in which we would express our conviction that ‘I feel it is I who raise my hand’, or ‘I experience a self’, or even ‘I have a feeling of doing’. When such things are uttered, we are already deep in theory. Indeed, common sense fails more dramatically here, since it is possible to identify a hand by pointing at it, but it is not possible to simply point at a self or a will.

I think it is relatively easy to discredit [INTROSPECTION 1] as it stands, simply by taking a second look and daring to question its results and assertions. When we are consciously focusing our gaze upon our hand and thinking of raising it, there will indeed be a ‘self-centred’ feeling of doing. But in normal cases when we are raising a hand, whether to grab something or ask a question or shake another hand, is there a corresponding feeling of deliberate doing – an experience, as Daniel Wegner (2002) would put it, of conscious will? Similarly, in the course of our everyday actions, are we usually actively deliberating and deciding about what to do before doing it? Are we usually *present* in our actions in the way that [INTROSPECTION 1] suggests we are – as consciously wanting, willing, doing selves? For my part, when I reconsider my daily dealings in the world, I think the answer to these questions should be no.

To look at a concrete case, let’s consider the action of taking a shower. Wegner uses this example in order to argue that an action cannot qualify as ‘truly *willed*’ (Wegner, 2002: 3) unless it was accompanied by an experience of conscious will:

‘If a person plans to take a shower, for example, and says that she intends to do it as she climbs into the water, spends fifteen minutes in there scrubbing up nicely, and then comes out reporting that she indeed seems to have had a shower but does not feel she had consciously willed it – who are we to say that she did will it? Consciously willing an action requires a feeling of doing [...], a kind of internal “oomph” that somehow certifies authentically that one has done the action. If she didn’t get that feeling about her showering, then there’s no way we could establish for sure whether she consciously willed it’ (*Ibid.*: 4).

Something appears to have gone wrong here: not just in the notion that it would be natural for us to experience an activity such as showering as consciously willed, and that something important would be missing in the absence of this experience – but in the underlying assumption that it ‘*usually seems that we consciously will our voluntary actions*’, even if ‘*this is an illusion*’ (*Ibid.*: chapter 1, subheading).¹⁰ On the basis of [INTROSPECTION 1], this assumption would indeed seem to be warranted – but is it really?

We could in fact describe the phenomenology of showering in wholly different terms. Speaking for myself, my voluntary showers are not usually preceded by a conscious and deliberate decision, let alone accompanied by an experience of will. Most of the time I don’t actively decide to take a shower – I just *do* it. If asked later what I had been doing, I would say ‘I took a shower’. If asked why, I might say something along the lines of ‘I felt like taking one’. If quizzed by a philosopher, I would indeed explicate that ‘I consciously decided to take a shower’, and that it was an act of will. But this does not mean that at the moment of showering itself, I had any experience of conscious will at all. On the contrary: as long as I’m not employing a version of [INTROSPECTION 1] in the course of the shower, it’s usually a rather passive experience, during which fragments of thoughts and sense input drift in and out of awareness. On the whole, there seems to be no question of a conscious will or decisive self – no ‘I’ at all – just the experience of showering.

This may also have been what Jean-Paul Sartre had in mind with his notion of the ‘transcendence of the ego’. Consider the following example, as worded by Jones & Fogelin (1997):

‘When I am intensely interested in what I am doing – say, in reading an exciting novel – I never think of myself as reading; I am fully occupied with the narrative. But if, after I have put the book aside, someone asks me what I have been doing, I reply without hesitation, ‘I was reading a book.’ Where does this knowledge come from? Careful introspection reveals that no ‘I’ was actually present in my consciousness while I was reading the book. Nevertheless I now know that at that time I was reading. Further, the ‘I’ that is so seldom present is always available, on call. This too is shown by introspection: I can at any time recall either *what* I experienced on a particular occasion in the past or the fact that it was *I* who experienced it’ (371).

Like the shower scenario, this example demonstrates that a different mode of ‘careful’ introspection, or introspection at *second sight*, can amount to a different kind of self-experience. For lack of a better term, let’s call this alternative mode of access to ourselves [INTROSPECTION 2]. Using this mode, it seems to be the case for most voluntary activities – at least, for

most of mine – that we can describe them in different terms from our customary self-reports.¹¹ In other words, a careful comparative [INTROSPECTION 2] can be wielded to produce a different self-experience than that which results from [INTROSPECTION 1]. This suggests, somewhat paradoxically, that our most commonsensical introspective self-descriptions ('I read a book', 'I wanted to read a book') do not necessarily correspond to our ordinary default self-experience. And that, in a nutshell, is the introspective argument against introspection: philosophers may appeal to introspection in support of a certain notion of self-experience – yet it is introspection itself that can show us that introspection can distort.

Foreground, Background

At this point it may be helpful to introduce the concepts of a foreground and background of experience. I suggest that we may speak of the *foreground* when we are consciously paying attention or thinking about something, and when we are consciously deliberating or deciding. It is here that we may encounter an inner voice, a conscious will, and a centred self. All the rest is in the *background*, which can perhaps be conceived as a dynamic, multi-faceted, often fuzzy web of semiconscious and possibly preconscious footage at the rim of our awareness. Here we may find half-digested thoughts, emotions, intuitions, perceptions, and even actions: things that are not interpreted by the foreground, but are nevertheless part of our experience, and can be stored for future use.¹²

It should be noted that this schematic distinction between a foreground and a background is *not* meant to correspond to the dichotomy of conscious versus unconscious or automatic behaviour, which is often presupposed by philosophers, psychologists and scientists alike. For instance, psychologists such as Bargh & Chartrand (1999) discuss an array of empirical evidence to argue that 'most of our day-to-day actions, motivations, judgments, and emotions are not the products of conscious choice and guidance', but are 'driven by automatic, nonconscious mental processes' – indeed, 'it appears impossible, from these findings, that conscious control could be up to the job' (464-5). Underlying their thesis is a clear-cut distinction between conscious processes on the one hand, which are associated with awareness,

intent, effort and control, and nonconscious processes or automatism on the other (463). To me, such a black-and-white distinction of conscious control versus nonconscious automatism seems deeply flawed, as it overlooks the wide grey zone of pre- and semiconscious aspects of experience.¹³ Consequently, it makes the primacy of the nonconscious seem counter-intuitive, as it stipulates that such processes pass outside the reach of our experience, awareness and agency.

This does not have to be the case. What the concept of a background may help us to grasp intuitively, is that even if an action is not actively, consciously *willed*, this does not mean it is not part of our awareness, our experience – of *who we are*. The boundaries between foreground and background should be visualised as always flowing, shifting, oscillating, fluctuating: whether continuously, as in James' concept of a 'river' or 'stream' of consciousness (1890/1950: 239), or discontinuously, 'constantly broken by detours – by blows – fissures – white noise' (Strawson, 2003: 356). The difference between the grounds is gradual, dynamic, and unstable: what is now in the foreground may merge into the background before we know it, and bits and pieces of the background may pop up into the foreground and evaporate again in the blink of an eye.¹⁴ Actions may flow from the background as well as the foreground, and this does not disqualify either kind from being rightly attributed to our (possibly retrospective) sense of agency.

Hence, I do not deny that the foreground is a real and important part of our experience. Indeed, it could be argued that our most essential and self-defining actions spring from the foreground – for instance, when we are facing difficult decisions or moral dilemmas, when we are trying to figure out what course of action would be best, or when we are consciously reflecting on our behaviour. But even if the foreground is a real and important *part* of our lives, this does not mean it is the *only* part. I believe experience can show us that most of the time, we are living in the background, although the foreground is always on call. In fact, the very moment we want to take a closer, more conscious look at things, the shift is made, and we are already in the foreground. And this is why we cannot trust the kind of active introspection some philosophers endorse. For the problem with any variety of [INTROSPECTION 1] is that it is a child of the

foreground, and speaks no other language than that of a conscious, willing self. Hence, whenever we introspect or ‘retrospect’, any experience that usually belongs to the background is drawn into the foreground, and interpreted in terms of a one-dimensional self. Of course nothing is wrong with this in principle: let all introspect as they please. It is not until philosophers start drawing conclusions that things get messy.

My thesis is that [INTROSPECTION 1] by nature only interprets the foreground, *even where there was none* – and therefore, that we need an alternative along the lines of [INTROSPECTION 2] to get in touch with the background, which makes up a more integral part of our everyday experience. Here we appear to meet a paradox. For if, as I have suggested, the experience of the self is mediated by the reflective act, how can we reflect on our usual experience without such mediation? How can we use introspection to get behind introspection? Perhaps we should accept this paradox, since we cannot unravel the Kantian knot: we cannot access our experience except through our experience, and we cannot see the self without looking at it. But my claims are more modest than that. My point is rather that different ways of introspecting lead to different (and sometimes incompatible) introspective results. Furthermore, I believe the kind of mediation that comes with [INTROSPECTION 1] can be avoided to some extent by employing the ‘method of stealth’ of [INTROSPECTION 2], as I have tried to do in the shower example.

This consists less in active introspection than in a more passive procedure of monitoring our daily experience, to catch us ‘in the act’ of consciousness. One could say that instead of switching on the light and shouting ‘freeze!’, we could try to sneak up on ourselves in the dark, and thus hope to catch a glimpse of ourselves before the limelight of introspection is on. We cannot completely avoid the focused, deliberate introspective gaze, for the light must go on at some point. When it does, the phenomenon as it was (an unreflected embeddedness in the background) disappears, as the background is irrevocably drawn into the focus of the foreground and loses its fuzzy, unfocused character. But traces may remain, may be remembered. If nothing else, [INTROSPECTION 2] may prove useful in moderating the results produced by [INTROSPECTION 1], and supplementing them with its own.¹⁵

Conclusion

Summing up, there is a major risk in trusting any kind of introspection at face value, and there lies a phenomenological fallacy in relying too easily upon ‘phenomenal facts’. At the same time, phenomenology cannot go without the first-person perspective: any attempt to formulate a third-person phenomenology retains an element of naivety, since it fails to recognise that third-person descriptions ultimately go back to a first-person stance. If phenomenology is to survive at all, some kind of introspection is required. The question is, of course, *what* kind – since the chosen ‘method’ of introspection may determine the results. In the context of this paper, [INTROSPECTION 2] was used to demonstrate the possible deficits of [INTROSPECTION 1]. The latter often boasts of possessing an unproblematic, commonsensical, *prima facie* lucidity, yet it remains to be seen on a case-by-case basis whether it has enough explanatory power to uphold its bold conclusions. This is why it is healthy at times to use a version of [INTROSPECTION 2], to bring us as close to usual experience as possible, and at least try to look at it from the background’s point of view.

What, then, does this imply for free will philosophers? In my view, several do’s and don’ts. First of all, do introspect. And second, do draw the background into it. But third, don’t universalise your introspection; and fourth, don’t automatically equate it with your usual experience. Fifth and finally, don’t draw too clear a line between the conscious and the nonconscious – for there may be a whole realm of experience in between. Perhaps, if these lessons are learned, we can question the persistent *reductive* conception of the will and the self, among proponents of mental causation and their critics alike, as something that remains once we have subtracted from our behaviour all nonconscious (environ)mental factors that have any causal influence upon it. The main question is often considered to be whether such a conception makes metaphysical sense, when the real question should be whether it makes *experiential* sense.

I do not believe this is how we usually experience our agency, nor our identity. We do not just identify with the beam of active, deliberating, focused consciousness that is most often projected onto the philosophical stage. We also identify with the ‘semi-automatic’ stream of thoughts, actions and impressions, which are not (fully) conscious to us, but are

nevertheless *there*: the many things in the background that give meaning and shape to the foreground, and deserve not to be overlooked. If the self consisted only in the truly conscious part, not much would remain of it: it would not contain, for instance, the sudden flashes of insight, inspiration, or creativity, which appear to rise from nowhere, yet are so essentially part of who we are. Nor the kind and cruel acts we do without thinking, and afterwards may cherish, or regret. Such things may not be part of the focal point of our experience, but they are part of its horizon.

Trying to catch these fragile fragments of experience may take us to the limits of language, since we will have to resist the temptation to speak of the background in terms of the foreground. As Searle noted (1983: 157): ‘The price we pay for deliberately going against ordinary language is metaphor, oxymoron, and outright neologism.’ But what we lose in clarity, we gain in lifelikeness: for if the talk of backgrounds seems fuzzy, whoever said that experience was not?

Mara van der Lugt (1986) studied Philosophy at Erasmus University Rotterdam, and obtained her master’s degree cum laude in 2010. For her second master in Early Modern Intellectual History she received the Prof. Bruins Prize for best research master student at Erasmus University in the year 2010. She is currently pursuing a PhD in History of Philosophy on the writings of the 17th-century French philosopher Pierre Bayle, at Corpus Christi College, University of Oxford.

‘The Foreground and Background of Consciousness’ was written for the mastercourse ‘Experimenting Ethics Away’ (of dr. Maureen Sie), taught by dr. Leon de Bruin.

Literature

- Bargh, J.A., & Chartrand, T.L. (1999) ‘The Unbearable Automaticity of Being’. In: *American Psychologist* 54, no. 7, 462-479.
- Bayne, T. (2006) ‘Phenomenology and the Feeling of Doing: Wegner on the Conscious Will’. In S. Pockett, W.P. Banks & S. Gallagher (eds.) *Does Consciousness Cause Behavior?* Cambridge, MA: MIT-Press, 169-186.
- Bayne, T. (2004) ‘Self-consciousness and the Unity of Consciousness’. In: *The Monist* 87, no. 2, 219-236.
- Bradley, R. D. (1958) ‘Free Will: Problem of Pseudo-Problem?’. In: *Australian Journal of Philosophy* 36, no. 1, 33-45.
- Brown, S.R. (2000) ‘Tip-of-the-Tongue Phenomena: An Introductory Phenomenological Analysis’. In: *Consciousness and Cognition* 9, 516-537.
- Dennett, D.C. (1991) *Consciousness Explained*. London: Penguin Books.
- Dennett, D.C. (1984) *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge, MA: MIT Press.
- Gertler, B. (2008) ‘Self-Knowledge’. In E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2008 Edition). Available from: plato.stanford.edu/archives/win2008/entries/self-knowledge [30 November 2011].
- Ginet, C. (1990) *On Action*. Cambridge: Cambridge University Press.
- Gurwitsch, A. (1964) *The Field of Consciousness*. Pittsburgh: Duquesne University Press.
- Hohwy, J. (2004) ‘The Experience of Mental Causation’ In: *Behavior and Philosophy* 32, 377-400.
- James, W. (1890/1950) *The principles of psychology*. New York: Dover Publications.
- Jones, W.T. & Fogelin, R.J. (1997) *The Twentieth Century to Quine and Derrida*. Fort Worth: Harbourt Brace College Publishers.
- Kane, R. (2005) *A Contemporary Introduction to Free Will*. New York/Oxford: Oxford University Press.
- Lewis, M. & Stachler, T. (2010) *Phenomenology: an introduction*. London/New York: Continuum International Publishing Group.
- Libet, B. (1999), ‘Do We Have Free Will?’ In: *Journal of Consciousness Studies* 6, no. 8-9, 47-57.

- Moore, G.E. (1962) *Philosophical Papers*. New York: Collier Books, 1962.
- Moran, R. (2001) *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton/Oxford: Princeton University Press.
- Nahmias, E. (2002) 'When consciousness matters: a critical review of Daniel Wegner's The illusion of conscious will'. In: *Philosophical Psychology* 15, no. 4, 2002.
- Nahmias, E., Morris, S., Nadelhoffer, T., & Turner, J. (2004) 'The Phenomenology of Free Will'. In: *Journal of Consciousness Studies* 11, no. 7-8, 162-79.
- O'Connor, T. (1995) 'Agent Causation'. In T. O'Connor (ed.) *Agents, Causes, and Events*. New York: Oxford University Press.
- Polanyi, M. (1983) *The Tacit Dimension*. Gloucester, Mass.: Peter Smith.
- Searle, J.R. (1983) *Intentionality, an essay in the philosophy of mind*. Cambridge: Cambridge University Press.
- Searle, J.R. (1984) *Minds, brains, and science*. Cambridge, MA: Harvard University Press.
- Smith, D.W. (2011) 'Phenomenology'. In E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition). Available from: plato.stanford.edu/archives/fall2011/entries/phenomenology [30 November 2011].
- Strawson, G. (2003) 'The Self'. In R. Martin & J. Barresi (eds.) *Personal Identity*. Oxford: Blackwell Publishing Ltd. 2003.
- Wegner, D.M. (2002) *The illusion of conscious will*. Cambridge, MA: MIT Press.
- Wittgenstein, L. (1951/ 2007) *Philosophical Investigations*. ed. G.E.M. Anscombe. Oxford: Blackwell Publishing Ltd.
- Wittgenstein, L. (1969/1975) *On Certainty*. ed. G.E.M. Anscombe/ G.H. von Wright, Oxford: Blackwell Publishing Ltd.
- Whitehead, A.N. (1947) *Essays in Science and Philosophy*. New York: Philosophical Library.

Notes

- 1 Quoted in Bradley (1958: 38).
- 2 Phenomenology can be defined as 'the study of structures of consciousness as experienced from the first-person point of view.' (Smith, 2008: unpaginated). 'Phenomenology does not attempt to speak about things, but only about the way they manifest themselves, and hence it tries to describe the nature of *appearance as such*.' (Lewis & Stahler, 2010: 1). In the context of this paper 'phenomenological' designates any attempt to describe consciousness and experience from the first-person perspective. (See section 2 below.)
- 3 E.g. claims about experience that pretend to apply to all humans, such as: 'Everyone experiences the world as unified', or: 'We all have an experience of conscious will.'
- 4 The free will debate may also be construed as a conflict between determinism and the demands of moral responsibility, in which case experiential factors can be (but are not always) left out. For a general overview of the free will debate, see for instance Kane (2005).
- 5 Compatibilists (who believe free will to be compatible with determinism) appear to be less committed to a 'self as source'- phenomenology, and even tend to 'describe the deliberative process more passively, with our decisions "flowing from" our desires and beliefs' (Nahmias *et al.*, 2004: 167) – yet they do sometimes imply the kind of mental causation phenomenology that is described by Hohwy above. Though I will here mainly question the self-phenomenology implied by libertarians (who believe determinism to be false) and their opponents, this is not to say compatibilists are not at all implicated in my criticism below, as it is directed at any kind of unquestioned introspection.
- 6 A version of this, as phrased by Gertler (2008: unpaginated): 'Since an appearance of a phenomenal quality and the reality which appears (the phenomenal quality itself) are one and the same, on this account, one can enjoy epistemically direct access to the phenomenal quality by attending to it.'
- 7 See for instance Moran (2001: 27-8): '[...] the problem of self-knowledge is not set by the fact that first-person reports are especially good or reliable, but primarily by the fact that they involve a distinctive mode of awareness, and that self-consciousness has specific *consequences* for the object of awareness.'
- 8 Note that in this context 'usually' does not necessarily imply 'most of the time': rather the 'usual' experience of the self signifies a default mode of consciousness prior to introspection. Though in practice it is possible we are in this default mode most of the time, this is not *necessarily* so.
- 9 Moore famously delivered a 'Proof of an External World' by means of raising one hand, then another, in order to show how an appeal to common sense can solve or dissolve philosophical dilemmas (Moore, 1962: 144-8). Hand-raising examples appear here and there in the free will debate, as in Bayne (2006: 176) and, interestingly, in the experimental context of Libet 1999, where 'the sudden flick of the wrist' is considered a typical act of will (50). See also Wittgenstein (1953/ 2001: no. 621): 'Let us not forget this: when "I raise my

arm”, my arm goes up. And the problem arises: what is left over if I subtract the fact that my arm goes up from the fact that I raise my arm?”

10 Note that we can detect two kinds of universalising tendencies in Wegner’s argument: not just the universalisation from ‘I’ to ‘we’ as discussed above, but also a universalisation from one or a few instances of introspection to a wide range of human experience: from ‘now’ to ‘most of the time’. Searle seems to do the same in claiming that we experience free will ‘all the time’ (1984: 88). Though my argument is directed against any failure to discriminate between experience *pre-* and *post-*introspection, this second kind of unwarranted universalisation gives it particular urgency.

11 Consider Dennett’s alternative phenomenology of decision-making: ‘[...] decisions can also be seen to be strangely out of our control. We have to wait to see how we are going to decide something, and when we do decide, our decision bubbles up to consciousness from we know not where. We do not witness it being *made*: we witness its *arrival*.’ (1984: 78).

12 This notion of a background to experience is not the same as the background to belief as conceptualised by Searle (1983), or the ‘tacit dimension’ of Polanyi (1983), though they may be related. Gurwitsch comes close, yet is too static in his division of the ‘field of consciousness’ into the ‘*theme*’ or focus of attention, the ‘*thematic field*’ or relevant background to the theme, and the irrelevant data in the ‘*margin*’ (1964: 4). More flexible is James’ concept of a ‘*psychic overtone, suffusion, or fringe*, to designate the influence of a faint brain-process upon our thought, as it makes it aware of relations and objects but dimly perceived’ (1950: 258). Whitehead also speaks of ‘the dim background of our conscious experience’ (1947: 122).

13 These include things of which you are not really conscious at the moment itself, though in retrospect you may realise they were in your awareness. For instance: a neighbour is playing some nice classical music. It is not until the music stops that you realise you were listening and enjoying it the whole time. It may have been in the background – but it was there. One might also think of actions that were not accompanied by a conscious intentional stance, though in retrospect one would appropriately ascribe an intention to them.

14 Tip-of-the-tongue phenomena are interesting in this context, as they may make us aware of an absence or background to experience (Brown, 2002); of ‘a gap that is intensely active’ (James, 1890/1950: 251).

15 There still seems to be a paradox in the fact that I too am drawing on my own introspective experience as a basis for my argument, and I can prove neither that I am correctly describing my experience, nor that this description would match the experience of others. Again, I accept this paradox, firstly because I believe it to be unavoidable in any phenomenological discussion, and secondly because I do not mean to propose an alternative universalistic account of the experience of conscious will. I am trying not to generalise, but to *de-generalise*: to question the results of one common kind of introspection by exposing them to a second, more comparative kind. Perhaps my use of introspection should likewise be questioned and reconsidered: in fact my argument would encourage precisely such a re-examination.



This work is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License. For more information, visit <http://creativecommons.org/licenses/by-nc/3.0/>