

The Rationality of Saliency

Sebastian Watzl

Abstract

Is it rational to pay attention to something just because it was salient, and you felt your attention drawn to it? I will argue that, generally, the answer is: yes. My arguments exploit an analogy according to which mental saliency stands to attention roughly like perceptual experience stands to belief. I provide three arguments that resemble arguments to the effect that an agent, other things being equal, is rationally entitled to believe what she appears to see. First, a pragmatic argument: we need to rely on saliency to act at all. Since relying on saliency is the best we can do, we cannot be rationally required not to rely on saliency. Second, a phenomenal argument. From a first-person perspective it makes sense to attend to what is most salient because we implicitly accept the authority of what our current situation demands of us regarding attention. This implicit authority shows up in how saliency manifests in consciousness. Third, a function argument that provides a third-person vindication relying on the claim that the function of the saliency system is to automatically direct the agent's attention to what most needs it in her current situation.

1. Introduction

Is it rational to pay attention to something just because it was salient, and you felt your attention drawn to it? I will argue that, generally, the answer is: yes.¹

One might think: No, saliency doesn't give you reasons and any aspect of agency that is not based on agent responding to relevant reasons is either irrational or outside the scope of rational evaluability; reasons can be good or bad (or one's actions badly based on them), but an assessment as rational can't be made of any 'action' that only has causal influences not resembling reasons at all. By contrast, I will argue that it is, in a specific sense, rational to just follow where the attention tends to flow. Drawing too tight a connection between reasons and rationality would miss that form of rationality.

I argue for this general conclusion about rational agency through an extended case study. The relevant instance of agency is attention, and the relevant instance of an influence on it is *mental saliency*. Mental saliency, as I am using the term, is a specific factor in the flow of the mind.² It is an aspect of an agent's psychology. To a first approximation (for more below), as I will use the term, something is salient to an agent at some time and to the degree to which its mental representation is consciously felt to automatically push, pull, draw on, or capture the agent's attention (I will use those terms interchangeably). My conclusion will be that it is, other things being equal, rational for an agent to attend to something merely because they feel their attention drawn to it. More precisely, I will argue for the following claim:

The Rationality Claim (R) In the absence of reasons to the contrary and in the absence of distorting factors, an agent is rationally entitled to focus more attention on what is more consciously salient to them at the time.

Mental saliency, on my understanding, thus primarily influences what occupies an agent's attention (in perception, thought, emotion, decision making, mind-wandering, and more). By

influencing the direction of our attention, salience then in turn shapes what an agent prioritizes at some moment,³ and what they select for action, belief or memory.⁴ Given the downstream effects of attention as a gateway to belief and preference formation, action, emotion, memory, and more, the effects of salience on our lives are arguably very significant indeed. The influence of mental salience is not confined to isolated moments in our lives but shapes all agency; for example, both when it is automatic and when it is controlled.

The rational reach of salience, as articulated by the rationality claim, concerns its influence on attention. The claim does not entail that, for example, it is always rational to *prefer* the most salient option, to *click* on the most salient item in a social media feed, to *believe* in the objective significance of the most salient patterns in one's social world, or to *lend most credence* to the most salient utterances. The rationality claim does, by contrast, entail that it is, in the absence of reasons to the contrary and distorting factors, rational to *have one's attention occupied by the choice* one feels one's attention drawn to, *pay more attention to more salient options*, to *look attentively at a salient item* on social media, to *focus attention on the most salient patterns in one's social world*, and to *pay more attention to more salient utterances*. How much of relatively more attention salience rationalizes and for how long will depend on how salient the relevant mental representation is and whether its salience persists in time.

The rationality claim, thus, is a claim about the rationality of a certain form and distribution of attention. Whether what is downstream of a rational pattern of attention will be rational will depend on what factors other than attention affect those downstream effects and those factors' rational standing and role.

2. The apparent irrationality of salience

The rationality claim is not trivially true. The claim that we have a rational entitlement to attend to what is salient to us is no more trivial than the claim that we have a rational entitlement to believe what we see. Attention like other activities *can* be performed for reasons. A night guard, for example, pays attention to what is happening around the building for the reason that it is part of their job. And a student pays close attention to what is in the textbook for the reason that it will serve in the exam. If we had a rational entitlement to do something only if it was based on sufficiently good reasons for it, this would also apply to attention.

While we *can* pay attention to something based on reasons, mental salience seems not a factor that provides such reasons and hence – for those in the grip of connecting rationality and reasons – a non-rational influence on attention. Consider the following three examples.

First, think of perceptual salience. The salience of a sudden noise disrupts your attention against what reason seems to demand: to keep focused on the paper you need to finish. Or think of the pop-up windows on your computer screen or advertisements in your social media newsfeed: they are visually salient, and you end up reading what they have to say – whether you see any reason for doing so, or not. Perceptual salience's influence on perceptual attention seems to be a brute causal influence.

Consider also decisional salience. The impact of salience on choice is discussed in the heuristics and biases literature and in the discussion of framing effects (Tversky and Kahneman 1974, Gigerenzer and Todd 1999). The mental salience of the choice situation an agent takes themselves to face and the mental salience of a subset of the options they consider, and their properties impacts how the agent ends up choosing. Decisional salience is often presented as a paradigm of an irrational influence.

Finally, think of discussions of “nudging” (Thaler and Sunstein 2009). Saliency nudges affect an agent’s actions by pulling their attention to options a “choice architect” deems better for the agent.⁵ Some argue that such nudging does not properly engage people as rational agents because it operates through saliency effects. Blumenthal-Barby and Burroughs (2012, fn 6), for example, suggest that such a nudge “exploits nonrational elements of psychological makeup.”⁶

Saliency, as these examples illustrate, thus may seem to be brute causal influence on us and ‘saliency biases’ cognitive quirks of limited agents. Our attention happens to be naturally drawn to some items, features, or options. In ideally *rational* agents the influence of the mental saliency current that pulls our mind now here and now there would be neutralized. My arguments for the rationality claim aim to undermine this position.

3. The central analogy and overview of the arguments

My arguments exploit an analogy:

The central analogy Mental saliency stands to attention roughly like perceptual experience stands to belief.

To get a feel for the arguments, let us compare two cases. Think of an agent *Eve* who happens to find herself in a particular situation. Here she is, in a specific place and time and with her specific mental characteristics. Eve has not (yet) investigated the quality of her mental faculties. Now take first this case:

Attentional Eve Eve feels her attention drawn to something. This thing catches her eye, that thought feels natural to follow. Through patterns of mental saliency, she feels the demands on her attention by the situation at hand.

Is Eve rationally allowed to trust, or to take an unquestioning attitude toward (Nguyen 2021), what their situation appears to tell her regarding what to pay attention to, or is she rationally required to subject it to critical evaluation? According to the Rationality Claim she is rationally entitled to follow where saliency leads her. But the apparent irrationality of saliency suggests that she has no such entitlement.

To make progress, I consider a parallel case.

Perceptual Eve Here is Eve again. She opens her eyes and experiences a certain world around her. There are colors and shapes, and various objects appear to be there in her environment. Eve’s situation appears to tell her something, in her perceptual experience, about what is present in her immediate environment.

Is Eve rationally allowed to trust, or to take an unquestioning attitude toward, what her situation appears to tell her, in her perceptual experience, regarding what there is in the environment, or is she rationally required to subject her visual experiences to critical evaluation? In other words, is an agent, other things being equal, rationally entitled to believe what she appears to see?

Many philosophers have, in various forms, and with various qualifications, answered the latter question regarding forming beliefs on the basis of perceptual experience in the affirmative.⁷ In general, they have argued, it is, other things being equal, rational to believe what is presented to us in perceptual experience. Such a view, the central analogy suggests, is

like the rationality claim, but about perception and belief. I will argue that if Perceptual Eve is rational, then so is Attentional Eve.

Starting from the central analogy, I present three arguments for the rationality claim. First, there is a *pragmatic* argument: we need to rely on salience to act at all. Since relying on salience is the best we can do, we cannot be rationally required not to rely on salience. Second, there is an *argument from the phenomenal force* of salience in our conscious experience. I will argue that we experience salience as an attention command issued by our situation. And that we experience the situation as having a certain, limited, form of authority regarding such attention commands. This experienced authority provides the agent with a rational right to follow what the situation commands regarding her distribution of attention. Third, there is an *argument from well-functioning*. I will argue that we have reason to think that the salience system that produces attention commands as its output is a competent system: a reliable capacity to direct attention automatically to what most needs the agent's attention in her specific situation. I will suggest that these arguments are mutually re-enforcing.

The central analogy helps to put the apparent irrationality of salience in its place. Salience biases relevantly resemble perceptual illusions. The fact that some perceptual experiences mislead does not undermine our rational right to rely on perceptual experience. Similarly, the fact that in *some* cases, salience experiences are biased or misleading does not undermine our general rational right to rely on them. We are rationally entitled to treat illusion and bias, from our first-person perspective, as the exception rather than the rule.

4. Salience and attention

My arguments for the normative conclusion of the rationality claim, as the central analogy already indicates, rely on descriptive claims about salience and its relation to attention.

The terms 'salience' is used in different but overlapping ways. The rationality claim is not true for all these ways of understanding 'salience'. My aim is to carve out, relying in part on Watzl (2017), a phenomenon that, I argue, deserves the term 'salience', that is a fruitful phenomenon to theorize about, and for which the rationality claim is true and interesting.

'Salience', in its original meaning, refers to a "jumping" or "leaping" character.⁸ The phenomenon that is our focus here concerns leaping into our attention. Specifically, my arguments defend the rationality claim regarding what Watzl (2017) calls 'psychological salience' (I will often say 'mental' salience). While we may speak of external items as salient, most fundamentally salience is a property of mental states. Those mental states are psychologically salient. The external items they are about may then be called "externally salient". They are salient in virtue of the mental salience of the mental states that represent them.

Psychologically salient mental states, according to Watzl (2017), are occurrent, play a distinctive role in the guidance of attention (*ibid.*, Ch. 6) and, in paradigm circumstances for which I will defend the rationality claim, manifest in a distinctive way in the agent's conscious experience (*ibid.* Ch. 10). Watzl's use of 'salience' is intended as an extension of how the term is used in much of cognitive psychology, where items are, for example, described as salient when they, when perceptually registered, tend to trigger a non-strategic (or automatic) attention shift (*cf.* Wright and Ward 2008). We can roughly summarize relevant notion of salience as follows.

Psychological Salience Psychologically salient representations are a dedicated and specific *input* to attention. They are what guides an agent's attention in the absence of the exercise of executive control. In paradigm cases, they have a distinctive manifestation in phenomenal consciousness.

What we need from Watzl's understanding of psychological salience are two types of claims.

First, there are claims about **cognitive architecture**: psychologically salient representations are functionally distinctive. We can call them salience representations. They are the output of a dedicated and largely modularized system for automatic attention guidance. This system has no access to all the rest of the agent's cognitive and motivational system (it is partially informationally encapsulated and not isotropic in the sense of Fodor 1983). We thus can't change "at will" the salience of pop-up windows. According to the cognitive architecture Watzl defends, salience representations are to some degree "motivationally penetrated" (Watzl 2017): salience synchronically is *contingent* on the agent's task (cf. Folk, Remington and Johnston 1992; also Ansorge, Horstmann and Scharlau 2010), and *diachronically dependent on previous reward* for detecting the relevant items (Anderson 2016), *task history*, and potentially other factors. Changes to what is salient to an agent, over time though follow patterns that resemble perceptual learning: they often require many repetitions and are not responsive to a deliberate assessment of what would now benefit the agent's attention. Salience representations, according to Watzl, are constitutively involved in all automatic (non-controlled) attention guidance. The pragmatic argument and the argument from well-functioning will rely on these claims about cognitive architecture.

Second, there are claims about **conscious experience**: psychologically salient representations, the output of the modularized system just described, are subject-level mental states. Watzl argues that while a large variety of sub-personal processes and a large variety of features in the agent's environment cause the agent to enter into the relevant subject level state, psychologically salient representations are not sub-personal states. The relevant states, indeed, are normally characterized by a specific conscious phenomenology: the agent feels a draw on her attention. Psychologically salient representations have, as Watzl says, a 'felt motivational impact'. They seem to compel the agent to do something, namely to attend to those items. Arguably there are some unconscious psychologically salient representations. The rationality claim will be defended only for psychologically salient representations that are phenomenally conscious. I will call those salience experiences. The phenomenal argument will rely on claims about conscious experience.

The two component claims just mentioned help to make sense of the central analogy: just like perceptual experience is the conscious output of a partially modularized input system to belief formation with a distinctive phenomenology and function (perceptual modules; cf Fodor 1983), so conscious psychological salience is the conscious output of a partially modularized input system to attention with its own distinctive phenomenology and function.

My aim in this paper is to argue for a normative claim about rationality from those descriptive claims about cognitive architecture and phenomenology. Insofar as the latter claims are correct, the rationality claim is correct as well. If they are incorrect, then the lesson to learn is conditional: like in the case of perception and rational belief, cognitive architecture and phenomenology have consequences for matters of rational action.

Before we move on, note that while the central analogy concerns *perceptual* experience and perceptual belief, the rationality claim does not just concern *perceptual* salience and perceptual attention. Salience experiences come in different forms:

- X is visually salient to A iff A has a visual salience experience or representation of X. The same goes for other sensory modalities. Attention gets drawn to perceptually represented items, features or patterns in our surroundings: e.g. a sound captures our attention; a voice feels familiar or pleasant and we begin listening to it, we look here and not there.
- Something is cognitively salient to A iff A has a cognitive salience experience of X. Attention gets drawn from one idea, topic, or thought to another. This is one way the wandering mind evolves in time when it hops from one currently salient idea to another, or gets stuck on a topic in rumination that keeps drawing the agent back in.
- Something is emotionally salient to A iff A has an emotional salience experience of X. we may feel drawn to what makes us angry, what we love, or what is disgusting to us (there is also an arguably more primitive valenced form of affective salience consisting of fast and primitive affective salience representations.⁹
- Something is decisionally salient to A iff a specific act of choosing, deliberating or deciding is psychologically salient to A. Here salience influences which acts of deliberation and choice occupy our attention, i.e., which options we consider at all, which first and most strongly, how seriously we take those options and how long they linger in the mind.

The rationality claim covers all forms of salience and attention for which the relevant architectural and phenomenal claims that describe their relationship are true.

5. Rational entitlement

One way to defend the rationality of salience would be to argue that salience presents an agent with a reason for attention. For example, one might try to argue that if something is salient to an agent then it subjectively seems important to her and that seeming is a reason to attend to it. This is not (as I mentioned at the beginning) how I will defend the rationality of salience.

Instead, I argue that we sometimes are rational in a form of agency that is not based on reasons, namely when we have a primitive form of entitlement to engage in the relevant action. My arguments are similar to arguments to the effect that we have a rational entitlement to use basic *belief* forming methods, even if we cannot provide any justification or reasons for the use of those methods. In those cases, we have an *epistemic* entitlement: an entitlement to hold certain beliefs and to use certain belief forming methods. In the salience case, we have a *practical* entitlement: an entitlement to engage in a certain form of agency (specifically, attending).^{10,11}

Rational entitlement, as I am using the term, is a species of rational permission. Entitlement has a dual in obligation: someone is rationally entitled to A iff they are not rationally obligated not to A.

One way to think of rational entitlement or permission is this: an agent has the relevant rational permission if she lives up to the normative standards that apply to her situation. Epistemic entitlement consists in meeting the normative standards constitutively associated with belief. Practical entitlement consists in meeting the normative standards constitutively

associated with decision making, choice and action. The three arguments below articulate why an agent that attends based on mental salience meets those practical normative standards.

Importantly my goal is only to argue that we *start* with a rational entitlement to rely on salience. We don't necessarily end there. Just like perceptual entitlement, an agent's practical permission to rely on salience might be defeated by either reason to distrust a specific salience experience (a type of practical rebuttal; cf Pollock 1974) or if the relevant salience experience is the result of an irrational process (in this case, the practical warrant may be 'undercut' (ibid)).¹² Both types of defeat probably come in degrees: a rational right to rely on salience is diminished to the degree to which there are countervailing considerations, or the relevant salience is the result of an irrational process.

Note also that having a practical rational permission for something does not entail that that thing is, for example, morally appropriate. Some patterns of salience, arguably, are harmful to others (e.g. Whiteley 2023). Whether all such cases are cases where the rationality of salience has been defeated is, in my view, an open question. It may also be that what is practically rational for the agent conflicts with what is morally right. Even an undefeated practical permission to attend does not deliver an all things considered right to attend in that way.

I will now present my three arguments for the rationality claim.

6. The Pragmatic Argument

The basic idea of the pragmatic argument exploits the central analogy in the following way: if we are to pursue basic cognitive projects such as forming beliefs about the external world at all, we must make use of perception; similarly, I argue, if we are to pursue the basic practical project of using attention in action, we must make use of salience in our attending. Since we must, we are rationally entitled to do so.¹³ Here is the pragmatic argument a bit more formally in outline (I'll assume the duality of rational requirement and rational permission):

- A1 Any agent needs to rely on salience as an input to attention to act at all.
- A2 If an agent needs to rely on salience as an input to attention, then she is not rationally required not to rely on salience as an input to attention.
- So, A3 Any agent is rationally entitled to rely on salience as an input to attention.
- A4 If an agent is rationally entitled to rely on salience as an input to attention, then she is rationally entitled to focus more attention on what is more consciously salient to them at the time.
- So, R The rationality claim: In the absence of reasons to the contrary and in the absence of distorting factors, an agent is rationally entitled to focus more attention on what is more consciously salient to them at the time. (from A3 and A4)

If the cognitive architecture claim above is correct, then any agent with that cognitive architecture has at each time two basic options for how to distribute her attention. Either she relies on the defaults set by psychological salience, or she controls how she guides her attention. But to control how she guides her attention in a rational way she must assess the reasons for attending in one way or the other. In order for a practical reason to be available to the agent, she must to some degree pay attention to that reason. Suppose, then, that we had no entitlement

to distribute our attention in accordance with the defaults set by psychological salience. We would thus be rationally obligated to distribute our attention in accordance with what we had most reason to attend to. But in order to search the space of possible objects of attention for those we have most reason to attend to we would have to already use our attention. It seems that we would never be able to get started in our attempt to fulfill our rational obligation.

Note that this problem is not specific to agents with limited information storage or reasoning capacities or to agents operating under time constraints. Any agent needs to act. Suppose that the agent had vast information or reasoning capacities and that they could operate at lightning speed. They would still be faced with the problem of how to determine which reasons to focus on in the concrete situation they find themselves in. Knowledge about which reasons are relevant when and where at some point will not help. It will just be another piece of information in the agent's storage. The agent would need another piece of information about whether that piece of information about which reasons are relevant is relevant for the agential problem at hand. Any agent needs to have a method that restricts the information drawn on for each specific decision problem to what is relevant to that problem. On pain of regress, the agent needs an entitlement to use a method that does not consist in having a justified belief about what is relevant when and where.¹⁴

The issue that any agent needs a system that delimits what is relevant for concrete decision making has, for example, been discussed in de Sousa's (1987) influential treatment of the rationality of the emotions: one of de Sousa's points is that the emotions need to assist reasoning because "[n]o logic determines *salience*: what to notice, what to attend to, what to inquire about" (p. 191). Since emotional valence is connected to salience, the emotions, according to de Sousa, are essentially involved in rational agency. Once we get salience experiences themselves clearly in view, we see that any of the emotion-specific issues, such as valence, are a red herring in this discussion: there is a pragmatic need for a salience system that restricts what we attend to in a specific situation, whether or not that system in a particular instance works through anything like valenced, evaluative or affective representations.

The pragmatic argument *is* specific to attention, though. Attention prioritizes and thereby coordinates information so as to allow for unified agency. It integrates the various representations available to the agent by prioritizing some parts over others (compare also Wu 2014, 2016, 2023a on attention as selection for action). Agents thus use their attention to select and prioritize for embodied and mental action. If this capacity were always employed on the basis of reason, it would be useless for exactly this essential function: if agents needed to use their reasoning capacities in order to decide what to pay attention to, they might as well employ those capacities directly in the service of the actions attention is employed to sub-serve. Attention, thus, couldn't perform one of its central functions if the agent always had to use her reasoning capacities to determine what to pay attention to.

One might object that such considerations may show that agents need *some* basic system that sets their focus of attention by default. They do not show, so the objection goes, that the salience system we in fact use is doing its job well. Practical entitlement requires meeting the normative standards associated with decision making, choice and action. Why think that responding to salience helps an agent meet those normative standards just because they cannot but help themselves to such a system? Agents may have an excuse to use salience to direct their focus of attention; but an excuse is not a rational entitlement, so the objection goes.

In response, consider that relying on salience is, for us, a *basic* method for practical decision making, choice and action. That is: to engage in practical decision making, choice and action

at all, we need to rely on salience. Given the claim about cognitive architecture, we either attend based on salience or based on cognitive control. The latter would enable us to assess the reasons for attending to something. But we cannot systematically use our capacity to attend to things based on what we have most reason to attend to, since that would get us stuck in the practical regress mentioned above. And we cannot dispense with the capacity for attention completely in our practical decision making, choice and action, since in that case we could not select a coherent course of action among all the possible behavioral paths compatible with what we believe and desire or judge to have rights or reason to (Wu 2014). So, either we are practically warranted to use the salience system, or we are not practically warranted in anything at all: since we cannot systematically search the space of possible targets of attention, we could not act on the basis of what we have most reason to do.

Our use of the salience system as a default for setting the focus of attention thus corresponds well to Reichenbach's (1949, 482) parable:¹⁵

A blind man who has lost his way in the mountains feels a trail with his stick. He does not know where the path will lead him, or whether it may take him so close to the edge of a precipice that he will be plunged into the abyss. Yet he follows the path, groping his way step by step; for if there is any possibility of getting out of the wilderness, it is by feeling his way along the path.

So, any rational agent must have a system that sets defaults for their attention. The salience system is what we have to feel our way. Using it is the best we can do to grope our way through the vast wilderness of potential objects of attention.

In the case of Reichenbach's argument aiming to establish our entitlement to rely on induction as a belief forming method, there is an influential objection (cf. Bonjour 1998). The objection is that appeal to practical indispensability of a method speaks only to practical entitlement, but not to any epistemic entitlement for the use of the method for belief formation. The practical indispensability of a method, so the objection goes, makes it no more likely that the beliefs it delivers are true. This objection does not apply in the present case. Here we have a pragmatic entitlement for the use of a practical capacity, and not for an epistemic one.

7. The Phenomenal Argument

I now turn to the phenomenal argument. We don't just happen to rely on salience in our attentive agency. From a first-person perspective it makes sense to do so. In this section I argue that this is so because we implicitly accept the authority of what our current situation demands of us regarding attention. This implicit authority shows up in how salience manifests in consciousness.

Let us start with a comparison to perception again. Perceptual experience has a characteristic phenomenology or phenomenal character (what it is like) such that it 'makes sense' from a first-person perspective to respond with belief. It presents the world outside to the agent. The way it presents the world to the agent in her own experience is the phenomenal content of the experience.¹⁶ Several philosophers have argued that it is, other things being equal, rational for the agent to take at face value what her experience tells her (there is no rational requirement to check the accuracy of her own experience).¹⁷ I defend a similar phenomenal argument regarding salience, i.e. an argument regarding how the manifestation of salience in consciousness provides *it* with a rational force. The phenomenological manifestation of

psychological salience is such that it makes first person sense to respond with attention. We are rationally entitled to take our salience experiences at face value.

Here is the argument in somewhat regimented form:

- B1 If something is consciously salient to an agent, then it has a characteristic manifestation in phenomenal consciousness (phenomenal salience).
- B2 If an agent takes that phenomenal salience at face value, she will respond by paying more attention to what is more consciously salient to her.
- B3 In the absence of reasons to the contrary and in the absence of distorting factors, an agent is rationally entitled to take phenomenal salience at face value.

So, R The rationality claim. (from B1-B3)

To assess the phenomenal argument, we need to consider whether and, if so, how salience manifests in phenomenal consciousness.

That psychological salience manifests in phenomenal consciousness is the cornerstone of the claim about conscious experience I mentioned in Section 4. The claim does not preclude unconscious mental salience, just like the existence of unconscious perception does not preclude that perception, when it is conscious, has a characteristic manifestation in consciousness. When understood in this way, the existence of a phenomenal manifestation of psychological salience should not be extremely controversial. It is a central element in the phenomenology of pop-out (Watzl 2017), distraction (which can be and sometimes is resisted with effort; cf. Watzl 2018), and mental affordances (McClelland 2020, Jorba 2020). To speak of a *characteristic* phenomenology of salience is meant to indicate that there is an interesting phenomenal property that is shared between different conscious cases of psychological salience. I don't provide much further argument for B1 here except for giving an account of a phenomenal property that is plausibly shared between the otherwise quite different cases of psychological salience.¹⁸

Does *how* salience manifests in consciousness let the idea that we can take it at 'face value' get a grip (and hence B2)?

The face value idea could be applied in direct parallel to perceptual experience if phenomenology of salience consisted in the appearance of a reason to attend to something or the appearance of a positive value in so attending. A salience experience would tell the agent that there is a reason to attend in some way (arguably in a non-conceptual format). An agent could then take the appearance of a reason for attending at face value by acting based on that reason.

Such an *appearance of reasons conception* of the phenomenology of salience implies that when an agent's attention is captured in perception by a fly that circles around her or drawn in mind-wandering from what thought to another, then she experiences something as speaking in favor of (a reason for) or something good (something valuable in) attending in those ways. An appearance of reasons view could, for example, draw on Chandra Sripada's recent work on valuationist structure of our motivational architecture (Sripada, unpublished), who explicitly applies the valuationist model to account for attentional capture (Fn 23).

In my view the appearance of reasons conception, unfortunately, is not a plausible account of the relevant phenomenology. Here is one reason why: suppose, generally, you experience an apparent reason to ϕ . Then, it seems, if you are reflective and understand that this is the character of your experience, you should come to believe that there appears (for you) to be a reason for ϕ -ing. But we don't typically come to believe that there appears to be a reason for attending to a distracting fly by reflecting on how it was (when salient) experienced by us. So, in my view, the appearance of reasons conception, whether or not anything like Sripada's account of our motivational architecture is correct, arguably misses an important phenomenal difference between the descriptive phenomenal content of perception and the phenomenal manifestation of salience.¹⁹

In my view, the best account the phenomenal manifestation of salience is as the experience of a demand or mandate issued by the something in your current situation (cf. Watzl 2017, Ch. 6+10. See Siegel 2014 for a related, though slightly different, account of soliciting affordances generally). On this view, the phenomenology of salience consists in an experience with an imperatival content. This view about salience, in this regard, is like the imperatival view of (the affective character of) pain defended by Martínez (2011) and Klein (2015). According to an imperatival view, salience experiences tell us to do something. They have an imperatival content of roughly the form: focus your attention on this! For example, focus your attention on this pop-up window! Focus your attention on this option! The imperatival view, as Watzl (2017) argues, is consistent with an intentionalist view of conscious experience; and explains how the phenomenology of salience is (objectionally) unified with descriptive representations (see below for my take on this). The imperatival view of the phenomenology of salience unlike the appearance of reason view above does not generate implausible reasons for belief.

The imperatival view, like the valuationist picture, is well suited to explain the impact of salience on an agent's motivational system. Phenomenal salience is what motivates us, as a matter of fact, most of the time to attend to what is salient. We experience motivation not from an associated desire or intention, but from the attention command itself. This is the same for ordinary commands. My son is motivated to wash his hands just by hearing my command (wash your hands, please!), not by the fact that this command generates a desire to wash his hands. Klein (2015) argues that the same is true the commands of the body we experience as pain: we are motivated to protect and deal with the relevant body part not because we have a desire to do so, but because it hurts (i.e. the painful experience). Similarly, here: we are motivated to attend to what is salient just because it is salient.

The imperatival view of phenomenal salience can make sense of a version of B2. We take a command at face value by doing as we are commanded to do.

Yet, can B3 be defended? Is it, other things being equal, rational to respond to attention commands in this way? It may seem like the answer is no: in general, we need a reason to respond to a command, e.g. that there are negative consequences for not doing so. It is not, in general, rational to always do as you are told.

Yet, I will now argue – drawing on Colin Klein's (2015) work on the imperatival content of pain experiences – that the imperatival view of salience experience can support B3. The way we experience those commands rationalizes, from the first-person perspective, acting on those commands.

We can start with the observation that not all commands are equally motivating. Sometimes we are motivated to follow a command and sometimes we are not. For example, if his

childhood buddy tells my kindergarten aged son to wash his hands before dinner, this is unlikely to motivate him to wash his hands. By contrast, if *I* command him to do so, then he is (normally and to some degree) motivated to wash his hands (I'm lucky, I know). So, what is the difference between my hand washing commands and those of my son's best buddy? Why are some of these commands motivating and others not? Klein 2015 provides an answer, which I will here put as follows:

Command and Authority An experienced command is intrinsically motivating to an agent A in all and only those cases where A to some degree explicitly or implicitly accepts the source of the command as authoritative regarding the relevant subject matter.

My son accepts my authority regarding hand washing, and that is why my commands motivate his hand washing behavior. So, in cases where the agent accepts the authority of a commander, she doesn't need an external reason to respond to the command. It is only once my son stops accepting my authority in this regard (I know the day will come, sigh), that my commands won't motivate him any longer; and I will need to give him reasons.

What we need for present purposes is this: the connection expressed in Command and Authority between being motivated by a command and experiencing the commander as having authority allows us to "read off" from the motivation whether someone experiences someone has having authority or not. By observing whose commands motivate my son and being careful to rule out external sources of motivation, you can tell that he implicitly accepts my authority. Similarly, I suggest, given that attention commands engage our motivational system by themselves, without a desire to do as we are commanded to do, or without us considering reasons to follow the command, it seems that there must be some authority we implicitly accept.

But who or what could this be? Who sends us attention commands? According to Klein (2015) pain is an experience of what *our own body* commands. We implicitly accept the authority of our own body, he argues. What, though, is it that we experience as the commander in the case of salience? There normally is, of course, no person who issues a command or we experience as mandating what to pay attention to. Whose authority could we be implicitly accepting?

My proposal is that we experience the relevant attention commands as issued *by our current situation*. This is, I take it, broadly aligned with how Siegel (2014, p. 58) talks about experienced mandates generally as what 'the immediate situation demands' as well as Merleau-Ponty's (2002, p. 1) talk of 'the solicitations ... of the current situation.' What then is 'our current situation'? The notion of the agent's current situation arguably does not allow for a precise characterization. Broadly, it is a part of the apparent world centered on the agent – the world as the agent currently takes it. Here with a bit more precision:

Situation An agent A's situation at some time t is the aspect of the world A takes to be given in A's total experience at t and which at t seems to instantiate the various properties A experiences the world as instantiating.

The agent's situation is what the agent currently meets. It's a portion of the world selected by her capacities and colored by emotions, activities, and current thoughts. An agent's situation includes what is currently perceived by her, what she consciously judges to be true, and her emotional take on things; what she experiences as offering opportunities for action and thought, and also what Jorba (2020) calls the cognitive horizon: items she is (dimly) cognitively aware

of as potential objects for thought, deliberation, or reasoning. An agent's situation includes her body and its features (though qua part of my situation my body has no special status). But an agent's situation does not include what she merely imagines (but takes to be false), what she merely supposes for the sake of argument, or what she wishes was the case. Situations can naturally be understood as having parts, which are some of the objects or features that taken together compose the situation.

The claim then is that attention commands are experienced as issued from a part of our current situation in this sense, i.e. by a part what we currently seem to be faced with.²⁰

So, to have a salience experience is to experience a command issued by a part of our situation to attend in some specific way. Given that we are generally motivated to do as our situation commands, Command and Authority entails that we implicitly accept the authority of our situation regarding what to pay attention to.

And now we get B3: it is rational to follow the commands of someone or something whose authority we accept or who we experience as authoritative. From our own perspective it has a say on these matters. And if from our perspective it has a say then it is, in the absence of reasons to the contrary and in the absence of distorting factors, rational to do as it says.

Here, then, is where we have arrived so far: given how mental salience manifests in phenomenal consciousness, it is rational to attend to what is salient. That is because this phenomenal manifestation shows that we implicitly accept the authority of our situation regarding what to attend to.

8. The Function Argument

At this point, another question naturally arises: the phenomenal argument might show that we experience the situation as authoritative, but is our implicit take on its authority warranted? Should we *really* trust our salience experiences? The *function argument* aims to provide a third-person vindication of what seems to be the case from the first-person perspective.

The function argument draws on an analogy Burge (2003) and Graham (2012) make about our epistemic entitlement to rely on perception: we have such an epistemic entitlement, they argue, because it is the function of perception to represent veridically. My argument is that we have a practical entitlement to attend to what is most salient since it is the function of salience to automatically direct attention to what is in the agent's best interest to attend to. Here is a representation of the argument.

- C1 Salience experiences are the output of a specific cognitive system, the salience system.
- C2 The salience system is individuated by having the function to guide the agent's attention to what best serves her interests in her current situation.
- So, C3 It is a priori that: if the salience system is functioning normally and is under normal conditions, then what is most salient is what best serves the agent's interests in her current situation. (from C1, C2)
- C4 If C3, then, in the absence of reasons to the contrary and in the absence of distorting factors, an agent is rationally entitled to focus more attention on what is more consciously salient to them at the time.

So, R The rationality claim. (from C3 and C4)

C1 is an aspect of the claim about cognitive architecture I suggested to assume in Section 4: there is a salience system that outputs salience experience as an input to attention. C2 then proposes that this salience system has a specific function: namely to direct the agent's attention to where it best serves the agent's interests. From this and assumptions about the relevant notion of 'function' it then follows that when the salience system is placed under normal conditions and is functioning normally that system will in fact direct the agent's attention to where it best serves the agent's interest. C4 then draws a connection between those claims about normal functioning and what is rational for the agent to do.

I should warn that a full defense of the function argument as an independent argument for the rationality claim would need a detailed defense of C4, which for reasons of space, I will not provide here. The function argument, in the present context, though, need not stand on its own but serves the purpose of providing a partial vindication or explanation of the first-person entitlement at issue in the last two sections.²¹ In this context, I hope it suffices to provide a more limited defense of the intuitive idea that an agent who attends in a way that normally best serves her interests is not acting irrationally.

With that caveat, let us get to the core question driving this section: why would agents come to have a system that, like the salience system, so substantially influences their agency in the absence of control by what she wants or intends? In other words, what is the function of the salience system? How did it, for example, evolve? Once we know what the salience is supposed to do, then we can consider whether on a specific occasion salience is well-functioning, is dysfunctional, or whether we are faced with a mismatch between a well-functioning salience system and a present environment (a well-functioning heart does not function well in a mismatched environment, like under water).

My interest specifically is in the *etiological* proper function (cf Millikan 1984, Neander 1995; Shea 2013) of the salience system.^{22,23} A type of system has the proper function F in this sense, when tokens of that system persisted and got replicated because they in the past did F. There will be normal conditions and a normal way of operating for functional systems in this sense: they are those conditions and those ways of operating that explain why that system persisted and was replicated. The heart has the function to pump blood, since hearts persisted and got replicated because they did so when operating in a specific way and under specific conditions.

So, what way of operating of the salience system might explain why it persisted in the cognitive architecture of organisms and got replicated from one generation to another?

The beginning of the answer, I believe, can be found in the fact that any substantial cognitive system is behaviourally decoupled from its environment in the following sense (see also Wu 2014, 2016): given its current perceptual input, and internal setup, many behavioural responses (including cognitive responses, such as what to think about and what to encode in memory) are possible for that cognitive agent. Therefore, any cognitive agent must prioritize among many possible objects to act upon. The pragmatic argument above showed that the system that makes such prioritizing possible must have an automatic component. This component is the salience system.

Such a system, to persist and get replicated, must, it seems, under normal conditions and when functioning well aid that flexible behavior. When would it do so?

I propose that salience systems persisted and got replicated in organisms since they guided the organism's attention to where its interests were served under those conditions. Here then is the proposal:

The interest function The etiological function of the salience system is to guide the agent's attention in ways that best serve the agent's interests in her current situation given that she does not exercise cognitive control.

What serves an agent's interest is what serves the agent's biological and social needs, their desires, wishes and preferences. Objects may serve an agent's interests in many ways. Many serve our interests whether or not we attend to them. The conditions of well-functioning for the salience system concerns interests that are served by *attending* to something at that time.

The claim that salience representations function to guide attention to what is in the agent's interest might be surprising: some of the paradigmatically discussed cases of how salience influences decision making, after all, are cases where what is salient is what is irrelevant to the agent's current task. In the cases most studied empirically, salience typically acts as a distractor from the agent's current task: it prevents her from performing that task as well as she otherwise would have by drawing attention away to a different item. But the study of task interference does not show that salience is always distracting. What is most salient to an agent is normally surprising (Horstmann 2015), task relevant (Yantis and Jonides 1990) and responsive to past reward (Anderson 2016). Standard distractor experiments are explicitly designed to bring apart salience and task. This does not hold everywhere though: when the red berries are salient on a fruit gathering expedition, their salience is not distracting. And even when salience is distracting and hence irrelevant to the agent's current task or project, this does not show – as I will argue – that what is salient is irrelevant *to the agent*: the distracting sound of an explosion, a crying child, or a bear in the forest are normally relevant to the agent in the sense that it is in the interest of the agent to attend to them, even when doing so is a distraction from their current task or project. The psychologically most illuminating cases of perceptual representation are ones where perception and belief come apart (consider the Müller-Lyer illusion). This does not, though, show that perception is unreliable or that the function of perception is not to represent veridically. Similarly, the existence of cases where what is salient is irrelevant does not show that it is normally not in the agent's interest to attend to what is most salient.

Why think the salience system has the interest function?

One way to get there is by generalizing from different cases of salience. I consider these to be the empirical considerations that indicate the function that individuates the salience system: it is a reliable competence to direct the agent's attention to what where attention currently best serves her interests.

Consider what kind of things tend to be salient to an agent: first, one of the most basic factors determining salience is surprise (Itti and Baldi 2009; for a detailed review and generalization see Horstmann 2015): what is unexpected given the agent's prior knowledge and belief tends to be salient. But such novelty is exactly what the agent has an interest to pay attention to in the absence of voluntary top-down control: what is new needs attention to be understood, to ask questions about it, and if irrelevant to the agent's project's it needs a quick glance of attention to then be rejected as irrelevant. Second, salience is not only dependent on features like surprise. Importantly for our purposes it is also dependent on what the agent is currently doing. In the literature, already mentioned above, this is described as the *contingent* capture of attention: what captures an agent's attention at a specific moment in time to a large degree

depends on what is relevant to the project's she is currently engaged in. In a famous study Yantis and Jonides (1990), for example, show that abrupt onsets only capture an agent's attention if they are potentially relevant for her task (see also e.g. Folk, Remington and Johnson 1992, Eimer and Kiss 2008, Ansorge, Horstmann and Scharlau, 2010). Similarly, for colors and shapes. Novel colors are more likely to capture an agent's attention if she is engaged in a color detection task than in a shape detection task. Such contingency of salience on the agent's projects is exactly what we would expect if the function of salience would be to serve the agent's interests. Finally, consider that salience is also value driven. If an agent has been rewarded for detecting a certain type of stimulus before then that type of stimulus will tend keep being salient for the agent and capture her attention (Anderson 2016 for an overview). Again, salience seems to be exactly sensitive to what, in general, is in the agent's interest to pay attention to.

So, when we consider both what, in the abstract, would make a salience system persist and replicate and what the salience system actually appears to do, when functioning normally and under normal conditions, we can discern that it appears to have the interest function.

If that is true, then salience system will guide the agent's attention to what serves the agent's interest to pay attention, when it is functioning normally and when it is in normal conditions. So, an agent who has a well-functioning salience system and operates under normal conditions will attend to what is in their interest to attend to.

I believe it is fairly intuitive that such an agent is not irrational for using that salience system. Still, let me give the gist of an *argument* for C4, the premise that establishes a link between the function of the salience system and the Rationality Claim. If a system with a certain function functions normally (no disfunction and under normal conditions), then it also satisfies a norm associated with that function (it can fail to satisfy that norm if it is not functioning normally or under abnormal conditions). Yet it is not generally the case that meeting a norm of proper functioning entails a practical entitlement or shows that an agent lives up to the standards that apply to action, decision making and choice. The natural norm of the heart is not a norm for the agent. But, I have argued, the salience system outputs subject level experiences. Those subject level states directly contribute to whole agent agency. Further, I have argued that any agent that acts at all and that has any cognitive system will have to have a salience system, otherwise it will not be able to steer a coherent behavioural path (see sec 0). The natural norm for salience systems therefore is a natural norm for all normal whole agent cognition and whole agent agency. It is for this reason that the conditions of normal functioning are privileged for the whole agent. It is from her point of view, and not just the point of view of one of her subsystems, that any conditions in which the salience system is malfunctioning will be accidents. Her system gets it wrong in this case, even though it gets it right normally and under normal conditions. An agent is rationally entitled to assume that, unless she has reason to the contrary, she has not been under abnormal conditions, since being in those conditions would be an accident from her own point of view. Therefore, in the absence of reasons to the contrary and in the absence of distorting factors, an agent is rationally entitled to use the deliverances of the salience system in her attending. But to do so just is to focus more attention on what is more consciously salient to them at the time.

As I said, for a full defense more would need to be said for C4. For now, just take this as a somewhat fuller articulation of the intuitive idea that if an agent's interest are normally served by attending in some way, she is rationally entitled to attend in that way.

This concludes the function argument. I have tried to establish that the rationality claim follows from a plausible assumption about the function of the system that produces salience experiences. This system functions to direct attention to where it is in the interest to the agent to attend. This, given the role of the salience system in our cognitive architecture, following parallel arguments about perception, delivers a rational right to rely on salience.

Conclusion

I have given three interlocking arguments for why we have practical entitlement to attend to what is most salient to us: the pragmatic argument, the phenomenal argument, and the function argument. Together they support the view that it is, against the traditional view of salience as arational or as an irrational bias, rational to attend to what is most salient to us in a specific situation. This rational entitlement can be defeated, but it is our starting point. It is rational to follow what the situation demands of us, and we can rationally trust those demands. It is not impossible that those commands lead us astray. But following them is the best we can do. We start with a practical entitlement for our attending. This is the most primitive and basic attention norm.

The result of my discussion is that Immanuel Kant (1775-76, p.62) is fundamentally wrong in his assessment that attention ‘... [is] only the useful for us, if [it is] under the free power of choice, so that involuntary attentiveness ... produce[s] much harm’. Attention is useful to us to a large degree *because* it is involuntary. The rationality of salience makes us rationally vulnerable. We stand open to influences on our mind outside our control. Salience can indeed be misleading; the salience system that produces them might be malfunctioning; the agent’s environment may be abnormal; or salience may have been coopted by irrational processes. Such cases often occur because either the environment or the agent’s salience system has been deliberately manipulated by an experimenter, another person, or a company out for its own profit. But this cognitive vulnerability is the result of an exercise of rationality. There is something morally objectionable about some forms of engineering other agents’ salience experiences. If the arguments in this paper are correct, it is not that it treats agents as irrational but that it exploits our rational capacities against us.²⁴

References

- Allport, D.A. (1987). Selection for Action. In H. Heuer and H.F. Sanders (eds), *Perspectives on Perception and Action* (pp. 395–419). Hillsdale, NJ: Erlbaum.
- Anderson, B. A. (2016). The attention habit: How reward learning shapes attentional selection. *Annals of the new York Academy of Sciences*, 1369(1), 24-39.
- Ansorge, U., Horstmann, G., & Scharlau, I. (2010). Top–down contingent attentional capture during feed-forward visual processing. *Acta psychologica*, 135(2), 123.
- Archer, S. (2022). Salience and what matters. In *Salience: A Philosophical Inquiry* (pp. 113-129). Routledge.
- Berridge, K. C. (2012). From prediction error to incentive salience: mesolimbic computation of reward motivation. *European Journal of Neuroscience*, 35(7), 1124-1143.

- Blumenthal-Barby, J. S., & Burroughs, H. (2012). Seeking better health care outcomes: The ethics of using the ‘nudge.’ *American Journal of Bioethics*, 12, 1–10.
- Boghossian, P. (2003). Blind Reasoning. *Proceedings of the Aristotelian Society*, Supplementary Volume, 77, 225–248.
- BonJour, L.. (1985). *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.
- Broome, J. (1991). *Weighing goods: Equality, uncertainty and time*. John Wiley & Sons.
- Burge, T. (2003). Perceptual entitlement. *Philosophy and Phenomenological research*, 67(3), 503-548.
- Burge, T. (2009). Primitive agency and natural norms. *Philosophy and Phenomenological Research*, 79(2), 251-278.
- Carroll, L. (1895). What the Tortoise Said to Achilles. *Mind*, 4, 278–280.
- Charland, L. C. (1995). Feeling and representing: Computational theory and the modularity of affect. *Synthese*, 105(3), 273-301.
- Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and social psychology bulletin*, 25(2), 215-224.
- De Sousa, R. (1990). *The rationality of emotion*. Mit Press.
- Jennings, C. D. (2020). *The attending mind*. Cambridge University Press.
- Duckworth, K. L., Bargh, J. A., Garcia, M., & Chaiken, S. (2002). The automatic evaluation of novel stimuli. *Psychological science*, 13(6), 513-519.
- Eimer, M., & Kiss, M. (2008). Involuntary attentional capture is determined by task set: Evidence from event-related brain potentials. *Journal of cognitive neuroscience*, 20(8), 1423-1433.
- Enoch, D., & Schechter, J. (2008). How are basic belief-forming methods justified?. *Philosophy and Phenomenological Research*, 76(3), 547-579.
- Fodor, J. A. (1983). *The modularity of mind*. The MIT press.
- Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human perception and performance*, 18(4), 1030.
- Gardiner, G. (2022). Attunement: On the Cognitive Virtues of Attention. In Mark Alfano, Jeroen de Ridder, & Colin Klein (eds.), *Social Virtue Epistemology*. Routledge
- Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford University Press, USA.

- Graham, P. J. (2012). Epistemic entitlement. *Noûs*, 46(3), 449-482.
- Hansen, P. G. (2016). The definition of nudge and libertarian paternalism: Does the hand fit the glove?. *European Journal of Risk Regulation*, 7(1), 155-174.
- Hausman, D. M., & Welch, B. (2010). Debate: To nudge or not to nudge. *Journal of Political Philosophy*, 18(1), 123-136.
- Heeney, M. (2019). Diachronic agency and practical entitlement. *European Journal of Philosophy*, 1-22
- Horstmann, G. (2015). The surprise–attention link: A review. *Annals of the New York Academy of Sciences*, 1339(1), 106-115.
- Huemer, M. (2007). Compassionate phenomenal conservatism. *Philosophy and Phenomenological Research*, 74(1), 30-55.
- Irving, Z. C. (2018). Attention norms in Siegel’s The Rationality of Perception. *Ratio* 32 (1):84-91.
- Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision research*, 49(10), 1295-1306.
- Jorba, M. (2020). Husserlian horizons, cognitive affordances and motivating reasons for action. *Phenomenology and the Cognitive Sciences*, 19(5), 847-868.
- Kant, I. 1775–76 [2013], *Anthropology* Friedländer (1775–1776). Translated by G. Felicitas Munzel, In: *Lectures on anthropology* (pp. 37–256). Cambridge University Press, Cambridge.
- Kiesewetter, B. (2017). *The normativity of rationality*. Oxford University Press.
- Klein, C. (2015). *What the body commands: The imperative theory of pain*. MIT Press.
- Lindenberg, S., & Papies, E. K. (2019). Two Kinds of Nudging and the Power of Cues: Shifting Salience of Alternatives and Shifting Salience of Goals. *International Review of Environmental and Resource Economics*, 13(3-4), 229-263.
- Luthra, Y. (2017). Self-trust and knowledge of action. *The Journal of Philosophy*, 114(9), 471-491.
- Martínez, M. (2011). Imperative content and the painfulness of pain. *Phenomenology and the Cognitive Sciences*, 10, 67-90.
- McClelland, T. (2020). The mental affordance hypothesis. *Mind*, 129(514), 401-427.
- Mole, C. (forthcoming). Emancipatory Attention. *Philosophers' Imprint*.
- Mole, C., Smithies, D., & Wu, W. (Eds.). (2011). *Attention: Philosophical and psychological essays*. Oxford University Press.

- Munton, J. (under contract). *Priority and Prejudice: The Epistemic Evaluation of Attention and Salience*. Under contract at Oxford University Press
- Millikan, R. G. (1984). *Language, thought, and other biological categories: New foundations for realism*. The MIT press.
- Neander, K. (1995). Misrepresenting and Malfunctioning. *Philosophical Studies*, 79(2): 109-141.
- Nguyen, C.T. (2022). Trust as an unquestioning attitude. *Oxford Studies in Epistemology* 7:214-244.
- Peacocke, C. (2004). *The realm of reason*. Oxford University Press.
- Pollock, J. L. (1974). *Knowledge and justification*. Princeton University Press.
- Pryor, J. (2000). The skeptic and the dogmatist. *Noûs*, 34(4), 517-549.
- Reichenbach, H (1949). *The Theory of Probability*, 2nd edition. Translated by E.H. Hutten and M. Reichenbach. Berkeley: University of California Press.
- Raz, J. (1999). *Practical reason and norms*. Oxford University Press
- Roth, A. S. (2017). Entitlement to reasons for action. *Oxford studies in agency and responsibility*, 4, 75-92
- Shea, N. (2013). Naturalising Representational Content. *Philosophy Compass*, 8(5): 496-509.
- Siegel, S. (2014). Affordances and the Contents of Perception. In: *Does perception have content?*, 39-76.
- Siegel, S. (2016). *The rationality of perception*. Oxford University Press.
- Smith, L., & Archer, A. (2020). Epistemic injustice and the attention economy. *Ethical Theory and Moral Practice*, 23(5), 777-795.
- Sripada, C. (unpublished). Human Agent Architecture Has aValuationist Structure, accessed Nov 16, 2023 at <https://www.dropbox.com/s/unez1ynzwoy7a0t/Sripada%20-%20Valuationism.pdf?dl=0> (through <https://sites.lsa.umich.edu/sripada/philosophy/>)
- Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.
- Todd, R. M., & Manaligod, M. G. (2018). Implicit guidance of attention: The priority state space framework. *Cortex*, 102, 121-138.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, 185(4157), 1124-1131.
- Watzl, S. (2017). *Structuring mind: The nature of attention and how it shapes consciousness*. Oxford University Press.
- Watzl, S. (2018). Consciousness and no self?. *Ratio*, 31(4), 363-375.

- Watzl, S. (2022). The Ethics of Attention: an argument and a framework. In Sophie Alice Archer (ed.), *Salience: A Philosophical Inquiry*. Routledge.
- Whiteley, E. (2023). A Woman First and a Philosopher Second: Relative Attentional Surplus on the Wrong Property, *Ethics* 133 (4):497-528.
- Wu, W. (2011). What is conscious attention?. *Philosophy and Phenomenological Research*, 82(1), 93-120.
- Wu, W. (2014). *Attention*, Oxford: Routledge
- Wu, W. (2016). Experts and deviants: The story of agentic control. *Philosophy and Phenomenological Research*, 93(1), 101-126.
- Wu, W. (2023a). *Movements of the mind: A theory of attention, intention and action*. Oxford University Press.
- Wu, W. (2023b). On Attention and Norms: An Opinionated Review of Recent Work. *Analysis*, anad056.
- Wright, R. D. and Ward, L. M. (2008). *Orienting of attention*. Oxford: Oxford University Press.
- Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: voluntary versus automatic allocation. *Journal of Experimental Psychology: Human perception and performance*, 16(1), 121.
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American psychologist*, 35(2), 151.

¹ This paper thus contributes to the growing literature on the normative evaluation of attention: cf. Archer 2022, Gardiner 2022, Irving 2018, Mole forthcoming, Munton under contract, Siegel 2022, Smith and Archer 2020, Watzl 2022, Whiteley 2023, Wu 2023b.

² cf. Watzl 2017 for this use of the term salience and how it guides attention. See below for more.

³ Watzl 2017, Jennings 2020

⁴ Wu 2014, 2016, 2023a. See also Allport 1987

⁵ cf. Lindenberg and Papie 2019 for an account of nudging in terms of “shifting salience”.

⁶ For similar assessments of nudging as interferences with rational decision making see Hansen (2016), Hausmann and Welch (2010).

⁷ e.g. Pryor (2000), Peacocke (2004), Huemer (2007).

⁸ The OED, for example, refers to a 1649 dictionary entry by Sir Thomas Browne (*Pseudodoxia Epidemica*) that speaks of what “... is observable in salient animals, and such as move by leaping; whereto the Locust is very well conformed”.

⁹ For example, in the incentive salience system, e.g. Berridge 2012, or in “primitive” affect, Zajonc 1980; see also Charland 1995; Chen and Bergh 1999; Duckworth et al. 2002.

¹⁰ Roth (2017), Luthra (2017) and Heeney (2019) also extend a specific epistemic notion of entitlement (cf. Burge 2003; Dretske 2000, Wright 2004) to the practical domain. The notion of a practical entitlement there is slightly different from my use of similar terminology.

¹¹ Some readers (namely those attracted to a form of Humeanism) might wonder whether it isn't after all *trivially* true that we have a rational entitlement rely on salience in decision making and action. As Hume famously said, isn't it the case that it would not be irrational for an agent to prefer the destruction of the whole world to the

scratching of her finger? The rationality claim would be trivialized by the following position: For any action or activity ϕ , in the absence of reasons to the contrary and in the absence of distorting factors, an agent is rationally entitled to ϕ . Call this the rationality by default view. If the rationality by default view were correct, then it would be rationally permissible for an agent to be influenced in her agency by any factor whatever, including, of course, salience. An analogue of the rationality by default view for belief would be implausible. The worry might, though, in the spirit of Hume be expressed as a dis-analogy between the case of belief formation and the case of action (the analogy that partially drives my argumentation): the right to believe must be earned. Beliefs we just happen to have, by default, are irrational (rationally impermissible). By contrast, so the worry, the right to perform an action or make a decision need not be earned at all. We have a rational right to perform any action, unless there is positive recommendation against it. Actions we just happen to perform, by default, are rational (rationally permissible). I believe that the rationality by default view is false. It is in tension with viewing the influence of salience on action as irrational (see Sec. 2). If actions are rational by default, then acting on the basis of salience should be rational and not irrational. Relatedly, it is also in tension with any view on which rationality requires having and responding to reasons (e.g. Raz 1999, Kiesewetter 2007). In my view, it is Broome (1991, Ch. 5) who shows most powerfully why it is false. He shows that there couldn't even be coherence requirements on action without rational requirements of indifference: there are some options or courses of action A and B, such that rationality requires that an agent is indifferent between A and B, which is inconsistent with the rationality by default view. Whether convinced by Broome's argument or something else, my discussion is restricted by the assumption that also practical rationality is not a completely unearned default. One aim of my discussion is to carve out middle ground between the rationality by default view and the reasons requirement. Rationality must be earned, but it can be earned by other things than having reasons.

¹² Susanna Siegel (2016, Ch 9), for example, argues that some salience distributions may have an inference-like etiology (etiologies that "mirror" inferential relations between beliefs). Those quasi-inferential etiologies may be irrational.

¹³ See e.g. Enoch and Schechter 2008 for an argument that targets basic belief forming methods:

¹⁴ This is the practical analogue to the problem of justifying basic methods of inference, like deduction (cf. e.g. Carroll 1895; Boghossian 2003; Enoch and Schechter 2008). Railton 2009 discusses a closely related regress without though specifically touching on the topic of attention and salience.

¹⁵ Cited from Enoch and Schechter 2008.

¹⁶ cf. Siegel 2011 for this way of thinking about the content of perceptual experience.

¹⁷ See e.g. Pryor (2000), Peacocke (2004), Huemer (2007). Note that my use of the phenomenal argument is compatible with views like Siegel (2016), on which an experience's etiology affects its total rational status.

¹⁸ Note that the relevant phenomenology concerns not attending to something (Wu 2011, Watzl 2017) but feeling drawn to attend to it (which leads to attending but is not identical to it).

¹⁹ Another view of the phenomenology of salience would be McClelland's (2020) proposal that some affordances for attention are potentiated. So, here the agent would, roughly, have a powerful (potentiated) experience of the possibility of attending in a particular way. It is as if the agent just feels strongly that attending in this way is about to happen. This potentiation account would arguably not support B2 (or its combination with B3). Just because an agent strongly feels that something is about to happen, does not entail that it is rational to do it by 'taking this feeling at face value' (whatever this would here mean). Yet, fortunately for the defender of the phenomenal argument such a view is also problematic. It explains the phenomenology in a way that disconnects the action-force of the relevant experience from both our overall motivational system, and our representational systems and is inconsistent with any intentionalist view of phenomenal character.

²⁰ One might object that my situation (or its parts) surely can't issue commands. It cannot speak, after all. In response, note that it doesn't actually matter whether the situation actually issues any commands. What is important is that it is *experienced* as issuing such commands. But further, while it is, of course, true that the situation does not literally speak it is also not true that perceptual experience literally says something to the agent. When we say that experience has representational content and, in this sense, "says" something, in the same sense experience can have anchored imperatival content and in this sense issues commands. There is no difficulty in representing (in consciousness) imperatival contents that are anchored to specific aspects of the world as experienced by the agent.

²¹ Compare Peacocke 2004, Ch. 3, for an argument that aims vindicate or explain perceptual entitlement.

²² Some of what follows draws on discussions with XXX and their unpublished work.

²³ Another possibility (Burge 2003) is, for the case of representational functions, to appeal to anti-individualism (externalism) about mental content. While this view doesn't have a direct analogue in our case (since the content

of salience representations is not descriptive, but imperatival), the case may be extended to cover our case as follows. According to anti-individualism, the content of a mental state depends on specific causal interactions between those mental states and what they represent, specifically those where the mental state was accurate. The parallel for imperatival contents would be that here the content depends on interactions where following the imperative provided an *apt* response (where aptness is for imperatival representations what accuracy is for descriptive representations). According to anti-individualism the imperatival content of states in the salience system is thus provided by a range of interactions where what those imperatival representations commanded was an apt response. But now the considerations above made plausible that following an attention command is an apt response only under those conditions under which what is to be attended is what serves the agent's practical interests. The conditions where what salient is what the agent interests to pay attention will thus enter into the individuation conditions of the states in the salience system. Note that Burge 2009 discusses primitive agency and practical norms. He doesn't draw consequences for rational action though.

²⁴ Acknowledgements: removed to preserve anonymity.