Penultimate draft. The final publication is available at http://link.springer.com/article/10.1007/s10670-016-9867-x. Read-only version: http://rdcu.be/nZs0.

Action Is Enabled by Systematic Misrepresentations

Wanja Wiese (wawiese@uni-mainz.de)

Abstract

According to *active inference* (which subsumes the framework of predictive processing), action is enabled by a top-down modulation of sensory signals. Computational models of this mechanism complement ideomotor theories of action representation. Such theories postulate common neural representations for action and perception, without specifying how action is enabled by such representations. In active inference, motor commands are replaced by proprioceptive predictions. In order to initiate action through such predictions, sensory prediction errors have to be attenuated. This paper argues that such top-down modulation involves systematic (but paradoxically beneficial) misrepresentations.

More specifically, the paper first argues for the following conditional claim. If active inference provides an accurate computational description of how action is enabled in the brain, then action is enabled by systematic misrepresentations. Furthermore, it is argued that an inference to the best explanation provides reason for believing the antecedent is true: Firstly, active inference provides a crucial extension to ideomotor theories. Secondly, active inference explains otherwise puzzling phenomena related to sensory attenuation, e.g. in force-matching or self-tickling paradigms. Taken together, these reasons support the claim that action is indeed enabled by systematic misrepresentations. The claim casts doubt on the assumption that representations are systematically beneficial to the extent that they are true: if the argument in this paper is sound, systematically beneficial *misrepresentations* may lie at the heart of our neural architecture.

Keywords: active inference, agency, ideomotor theories, misrepresentation, predictive processing, sensory attenuation.

1 Introduction

It may be intuitive to believe that representations are systematically beneficial for an agent to the extent that they are true. Although some misrepresentations may be harmless, and sometimes false assumptions may coincidentally lead to true beliefs, it seems strange to assume that misrepresentations could be systematically beneficial. There is, however, a growing number of authors who put forward the view that some ways of being systematically wrong are in fact useful (e.g., positive illusions such as optimism bias[1]). In section 2, I will provide

Penultimate draft. The final publication is available at http://link.springer.com/article/10.1007/s10670-016-9867-x. Read-only version: http://rdcu.be/nZs0.

a brief overview of possible cases of beneficial misrepresentation and constrain the set of misrepresentations relevant for this paper. After that, in section 3, I will review a prominent theory of action representation, according to which neural representations related to action and perception share a common code. This theory leaves a crucial aspect unspecified, namely how action can be initiated in the absence of classical motor commands. As I will show in section 4, active inference (which subsumes predictive processing) promises to solve this puzzle. Furthermore, the computational models provided by active inference involve systematic misrepresentations that are crucial to bring about action (section 5). In section 6, I will argue that active inference provides true descriptions of the computational mechanisms underlying action initiation. Finally, I will address some objections (section 7).

2 Varieties of misrepresentation worth wanting

In their seminal BBS target paper, McKay and Dennett (2009) explore the possibility that some *false* beliefs may have been systematically adaptive in our evolutionary past. How is this possible? And what could it mean that misbeliefs have been systematically adaptive? The authors approach this question by distinguishing irrelevant model cases of misrepresentation from more relevant ones. Four types of misrepresentation they distinguish are (1) forgivable design flaws, (2) culpable design flaws, (3) by-products, and (4) design features. Of these, only the fourth type (misbeliefs conceived as design features) qualifies as relevant, according to McKay and Dennett. The reason is the following:

> Such misbeliefs, unlike occasional lucky falsehoods, would have been systematically adaptive in the evolutionary past. Such misbeliefs, furthermore, would not be reducible to judicious – but doxastically noncommittal – action policies. Finally, such misbeliefs would have been adaptive in themselves, constituting more than mere by-products of adaptively biased misbelief-producing systems. (McKay and Dennett, 2009, p. 493)

Let us consider in slightly more detail how such misrepresentations differ from other types of misrepresentation. This will clarify what, for the purposes of this paper, qualifies as an interesting case of beneficial misrepresentation.

(1) Misrepresentations can result from certain design limitations (cf. McKay and Dennett, 2009, p. 494). An example of a *forgivable* design limitation is a very cheap watch that loses ten seconds per day. That can be good

enough, if you don't need an extremely precise time specification. The resulting misrepresentation may be forgivable—but it is not beneficial.

(2) Worse are misrepresentations resulting from *culpable* design flaws. Suppose the watch loses ten minutes per day. That would make the watch more or less useless for most purposes. In general, misrepresentations resulting from design flaws may or may not be tolerable, but they are not beneficial.

(3) Misrepresentations can also be by-products of beneficial functions. As an example, McKay and Dennett (cf. 2009, p. 495) consider a rear-view mirror. Objects in a rear-view mirror are closer than they appear. The reason for this is that the mirror is convex, in order to provide a wider field of view, which is beneficial. That objects appear to be farther away, however, is not beneficial, it is just a tolerable by-product.

(4) Of most interest are misrepresentations that can be considered as design features. Candidates for such misrepresentations are, according to McKay and Dennett, positive illusions. For instance, there is evidence that most parents have an overly positive image of their children. As McKay and Dennett put it:

> Wenger and Fowers (2008) have recently provided systematic evidence of positive illusions in parenting. Most participants in their study rated their own children as possessing more positive (86%) and fewer negative (82%) attributes than the average child. This better than-average effect, moreover, was a significant predictor of general parenting satisfaction. (McKay and Dennett, 2009, p. 506)

Believing that your children are nicer, smarter, etc., than they actually are, may turn you into better parents. This is therefore a candidate for a beneficial misrepresentation (other possible examples are discussed in Mendelovici, 2013; Trivers, 2000; Zehetleitner and Schönbrodt, 2015).

Whether or not this is a convincing example will be of no concern here. The purpose is only to illustrate what I mean by a systematically beneficial misrepresentation (SBM) in this paper. An SBM is (cf. again McKay and Dennett, 2009, p. 493):

1. not just acceptable, but useful for the agent,
2. not just a lucky falsehood,
3. not just a by-product.

The first condition emphasizes that SBMs do not simply result from a trade-off between speed and accuracy, or between metabolic cost and accuracy. Such trade-offs can be achieved by *fast and frugal heuristics* (cf. Gigerenzer et al., 1999). Heuristics can lead to cognitive biases, and therefore to systematic misrepresentations (cf. Tversky and Kahneman, 1974, p. 1124). At the same time, heuristics are beneficial, because they are efficient (saving time and computational resources). But we should not conclude from this that heuristics involve SBMs. As McKay and Dennett (2009, p. 497 f.) suggest, misrepresentations resulting from heuristics are best conceived as design flaws. Heuristics are beneficial, because they (usually) lead to accurate judgments. The fact that they are efficient (in terms of time and computational resources) is a feature; the fact that they can lead to systematic misrepresentations is a (tolerable) design flaw. Such design flaws are acceptable, but not useful for the agent.

Note that McKay and Dennett (2009) call misbeliefs beneficial when they (or the mechanisms producing them) have been adaptive in our evolutionary past. My aim is not to show that certain misrepresentations have been adaptive, but that they fulfill a useful function *now*.

In the rest of the paper, I will argue that certain peripheral sensory precision estimates in computational models in predictive processing (PP) are SBMs in the sense specified above: (1) they are useful, because they enable action; (2) they are not lucky falsehoods, because they enable action reliably; (3) they are no mere by-products, because they themselves fulfill a functional role that is necessary for action to be possible. Furthermore, I will argue that there are good reasons to believe that computational models in PP provide true descriptions: on the one hand, they explain otherwise puzzling perceptual phenomena; on the other, they offer crucial complements to prominent theories of action, most notably *ideomotor* (or *common coding*) *theories*, to which I will turn next.

3 Ideomotor theories of action

How does the brain represent perceptual and motor events? According to common coding accounts (cf. Prinz, 1990), the same neural structures become active when subjects perceive and when they act. As Bernhard Hommel puts it (in the context of his *theory of event coding*): “the basic units of both perception and action [...] are *sensorimotor* entities, in the sense that they are activated by sensory input (=perception) and controlling motor output (=action).” (Hommel, 2015, p. 2). This contrasts with the “classical sandwich” view of perception and action (as Susan Hurley, 1998, has called it), according to which perception and action are completely distinct processes that do not directly influence each other, but only mediated through cognition.

Penultimate draft. The final publication is available at http://link.springer.com/article/10.1007/s10670-016-9867-x. Read-only version: http://rdcu.be/nZs0.

Interestingly, already William James (1890) and Hermann Lotze (1852, pp. 313 f.) saw the neural activity underlying action and perception as inherently intertwined.[2] In James' view, "the distinction of sensory and motor cells has no fundamental significance." (James, 1890, p. 581). This is a core assumption of his *ideomotor* theory, according to which there are no purely sensory or purely motor "ideas", only ideas that activate (and are activated by) movements: "or, to express it again in psychic terms, *the idea of the movement M's sensory effects will have become an immediately antecedent condition to the production of the movement itself.*" (James, 1890, p. 586). A particularly elegant aspect of such a view is that it does not have to posit any motor commands in a strict sense. The reason for this is, as Hommel puts it, that "the agent only needs to 'think of' the representation of a wanted action effect to activate the motor pattern needed to produce it" (Hommel, 2015, p. 2).

However, this raises a puzzle: if merely thinking of action effects activates motor patterns, does this mean that it is also *sufficient* to initiate the corresponding action? Clearly, this cannot be the case, because we would not be able to imagine movements (without actually carrying them out) if it were true. So how *do* sensorimotor representations lead to action (given that their mere activation is not sufficient)? Ideomotor theories remain silent on this point. In fact, in an introduction to the *theory of event coding* (TEC), Hommel, Müsseler, Aschersleben, and Prinz explicitly restrict the scope of the theory:

> TEC does not consider the complex machinery of the "early" sensory processes that lead to them. Conversely, as regards action, the focus is on "early" cognitive antecedents of action that stand for, or represent, certain features of events that are to be generated in the environment (= actions). TEC does not consider the complex machinery of the "late" motor processes that subserve their realization (i.e., the control and coordination of movements). Thus, TEC is meant to provide a framework for understanding linkages between (late) perception and (early) action, or action planning. (Hommel et al., 2001, p. 849)

TEC and other common coding theories are thus primarily theories of the *representation* of perceptual and motor events, and they claim that there are shared neural structures representing both types of events at the same time (i.e., the represented events are *sensorimotor* events). However, it would be surprising if these representations did not play a significant role in the initiation of action. And this is clearly presupposed by proponents of ideomotor theories, as the following statement (already quoted above) shows: "the basic units of both perception and action [...] are *sensorimotor* entities, in the sense that they are activated by sensory input (=perception) and **controlling** motor output (=action).” (Hommel, 2015, p. 2; bold emphasis added).

So how does action come about? In fact, there are at least two problems that need to be addressed, since representations of sensorimotor events are also activated by perception (so “thinking of” an action effect can in principle be undermined by perceiving that the desired effect is not currently taking place):

1. How do sensorimotor representations lead to action (when the subject intends to act), and how is action inhibited (when the subject merely imagines or observes movements)?
2. How can perceptual input be prevented from interfering with sensorimotor representations?

Both of these problems are addressed in *active inference*, which subsumes the framework of predictive processing (PP).[3] I will explain the basic tenets of PP (and active inference) in the following section. After that, I will argue that action is enabled by systematic misrepresentations in this framework.

4 Predictive processing and active inference

The framework of predictive processing (PP) depicts neural activity as a hierarchically distributed process of minimizing precision-weighted prediction error (cf. Clark, 2013, 2016; Hohwy, 2013). Three features are particularly important: (1) A substantial amount of neural activity can be interpreted as *prediction*; (2) predictions lead to prediction errors that are weighted by *precision estimates* (which represent the confidence placed in prediction errors where low variance or reliable prediction errors have high precision); (3) this process takes place in a hierarchy.

(1) The brain does not passively wait for incoming sensory signals, but anticipates signals on the basis of internal estimates. These estimates are formed at all levels of a hierarchy, and capture, as we move from the bottom to the top of the hierarchy, “increasingly complex and more abstract features [...] in the processing levels furthest removed from the raw sensory input.” (Clark, 2013, p. 200). The estimates are generated using a *generative model*, which is a probabilistic model of how sensory events are caused by external events (formally, a generative model is typically given in the form of a *likelihood* and a *prior* distribution, cf. Friston, 2010, p. 129). At each level, predictions of estimates at lower levels are computed, which yield prediction errors that are used to update and correct estimates.

(2) Crucially, prediction errors are weighted by estimates of precision. As an illustration, consider the rubber-hand illusion (cf. Botvinick and Cohen, 1998). In this illusion, subjects visually perceive a rubber hand that is being stroked by the experimenter. At the same time, they feel their real hand being stroked, without seeing it (their real hand is concealed from view, but also being stroked by the experimenter). When both stroking events occur synchronously, many subjects report (after a while) that it feels as if their real hand was located where they see the rubber hand.

We can interpret the illusion as the result of a conflict between vision and proprioception. This conflict is dissolved in favor of vision, because visual estimates are expected to be more precise than proprioceptive estimates. In other words, given a prior assumption that synchronous events (here: the stroking of the real and the artificial hand) have a common cause (cf. Hohwy, 2013, p. 105), visual estimates of where the stroking events are happening yield a prediction error when compared to proprioceptive estimates. If visual estimates have higher associated precisions, the resulting prediction error will have a stronger influence on proprioceptive estimates. When the illusion kicks in, proprioceptive prediction errors are attenuated and the estimate of the hand's location is allowed to drift towards the location of the rubber hand.

(3) Precision estimates are thus crucial to resolve conflict between modalities, but they actually play a far more prominent role, since conflicts not only occur between estimates related to different sense modalities. On the contrary, conflicts arise constantly between estimates at all parts of the hierarchy, and the computation and minimization of prediction error serves to dissolve these conflicts—with the help of precision estimates (cf. Clark, 2015, p. 8). As we will see below, optimizing precision estimates has been interpreted in terms of attentional processing (cf. Friston & Stephan, 2007; Feldman & Friston, 2010).

So the most general computational description of neural activity, according to PP, is simply *prediction error minimization* (in fact, this is the term Hohwy, 2013, uses to denote the framework). Starting from this very general principle, we can identify two strategies to minimize prediction error: perceptual inference and active inference (cf. Friston et al., 2012a).[4]

Put bluntly, perceptual inference is a process in which the generative model is changed in such a way that it yields more accurate estimates of external states. Active inference is a process in which also changes in the world

Penultimate draft. The final publication is available at http://link.springer.com/article/10.1007/s10670-016-9867-x. Read-only version: http://rdcu.be/nZs0.

are brought about that have reliable sensory effects (so predictions become more reliable, too, cf. Hohwy, 2013, pp. 79-81). The causal antecedents of these changes are estimates of hidden states in the world (and of their sensory effects). If I am intending to move my arm, my brain will predict that my arm is moving, and this prediction has implications for sensory estimates (e.g., visual and proprioceptive estimates), which will then be fulfilled by action. Active inference thus shares the assumption made by common coding (or ideomotor) theories that motor events are represented in terms of the effects they bring about:

> In this picture of the brain, neurons represent both cause and consequence: They encode conditional expectations about hidden states in the world causing sensory data, while at the same time causing those states vicariously through action. [...] In short, active inference induces a circular causality that destroys conventional distinctions between sensory (consequence) and motor (cause) representations. This means that optimizing representations corresponds to perception or intention, i.e. forming percepts or intents. (Friston et al., 2011, p. 138)

In other words: there are sensorimotor representations that cause motor events, but the same representations can also be caused by perceptual events. So far, this does not extend the theories mentioned above. In particular, we have yet to address the two questions posed in the previous section. This is where peripheral sensory precision estimates (PSPEs) become essential.

Note that sensorimotor estimates (“ideas”) corresponding to intended (but not yet performed) movements are in conflict with perceptual input. When I am intending to move my arm, the brain will predict the sensory effects this movement will have. For instance, if I am looking at a pencil in front of me on the table and intend to grasp it, my brain predicts visual and proprioceptive signals according to which my arm is being stretched and the hand is brought closer to the pencil. These predicted signals are in conflict with the sensory signals the brain is actually receiving (because my hand is not moving yet). The result is a prediction error.

There are two ways in which this error can be eliminated, and estimates of precision play a crucial role (recall from above that precision estimates determine how much influence prediction errors have on updates of estimates). Either the estimate according to which the hand is moving is revised (to veridically represent the absence of movement); or sensory signals are ignored, so the estimate can be sustained (and action is possible).

Loosely speaking, this entails a suspension of disbelief in the evidence for an absence of movement. From the point of view of PP, this temporary suspension can be thought of in terms of attention (see below). In other words, we attend away from evidence that we are not moving to enable our predictions to be fulfilled. Strictly speaking, this is of course not a process we as persons are carrying out, but a (usually unconscious) process taking place in our brains.

Which of the two options is realized depends on the relative precision of sensory prediction errors. If the relative precision of sensory prediction error is high, the estimate according to which the hand is moving will be corrected (which means that the predictions derived from it change, as well; this quashes the prediction error). On the other hand, if the relative precision of sensory prediction error is low, the sensorimotor estimate will be sustained and movements are generated through motor reflex arcs (cf. Shipp et al., 2013, p. 712). In this case, predictions derived from sensorimotor estimates do not change, but prediction error is eliminated because the predicted changes in the world have been brought about (and hence sensory signals will conform to the prediction that the hand is moving).

So action is only possible if sensory prediction errors are attenuated. This is the mechanism that prevents perceptual input from interfering with sensorimotor representations (motor intentions). But how are sensorimotor representations "translated" into motor trajectories? In active inference, this happens automatically, because sensorimotor estimates (which code movements in terms of their sensory effects) also yield proprioceptive predictions (in particular, about the current state of the muscles):

> A proprioceptive prediction error can be generated at the level of the spinal cord by the comparison of proprioceptive predictions (from motor cortex) and proprioceptive input. Sources of proprioceptive input include muscle spindles (via Ia and II afferents), Golgi tendon organs (via Ib afferents), and articular and cutaneous receptors. The prediction error can then activate the motor neuron to contract the muscle in which the spindles—or other receptors—are sited: this is the classical reflex arc [...]. In short, peripheral proprioceptive prediction errors are (or become) motor commands. (Adams et al., 2013, p. 614)

Crucially, in order for peripheral proprioceptive prediction errors to function as motor commands,[5] their associated precision must be high enough:

Penultimate draft. The final publication is available at http://link.springer.com/article/10.1007/s10670-016-9867-x. Read-only version: http://rdcu.be/nZs0.

> Active inference (action) occurs when peripheral proprioceptive prediction errors are afforded high precision, facilitating the spinal reflex (Adams et al., 2013; Friston et al., 2011); conversely, the precision can be attenuated (cf. sensory attenuation) during action observation (Friston et al., 2011). This attenuation does not affect action observation, because perceptual inference is driven by visual prediction errors. (Shipp et al., 2013, p. 712; citation style adapted)

This is the mechanism that prevents action when we merely imagine or observe movements (with no intention to actually perform or mimic the movements). Active inference thus provides computational descriptions of the mechanisms that enable (or inhibit) motor sequences. It thereby answers the two questions left open by ideomotor theories. Since the theory can be difficult to grasp at first, let me summarize the main ideas once again.

Figure 1 illustrates a peripheral part of the predictive processing hierarchy. Based on sensorimotor estimates (top), sensory (including proprioceptive) predictions are generated. When a subject intends to move, say, her hand, these predictions will correspond to sensory signals that *would be obtained* if the hand was actually moving. Hence, when being compared to *actually* obtained sensory signals, they generate large prediction errors. These errors are fed back to sensorimotor estimates, resulting in a correction—*unless* the errors are attenuated through top-down modulation. Intended movements can only be initiated if these errors are down-weighted (by modulating their associated precision estimates) prior to movement onset. When such top-down modulation takes place, sensorimotor estimates (according to which, say, the hand is moving) can be sustained and the resulting proprioceptive prediction error can be quashed through classical reflex arcs (which means that the intended movement is carried out).

The essential aspect of this account is that peripheral sensory precision estimates (that function as weights for prediction errors) must be turned down during action (otherwise action is inhibited). This, however, means

Penultimate draft. The final publication is available at http://link.springer.com/article/10.1007/s10670-016-9867-x. Read-only version: http://rdcu.be/nZs0.

that the precisions of peripheral sensory signals are systematically *underestimated* prior to

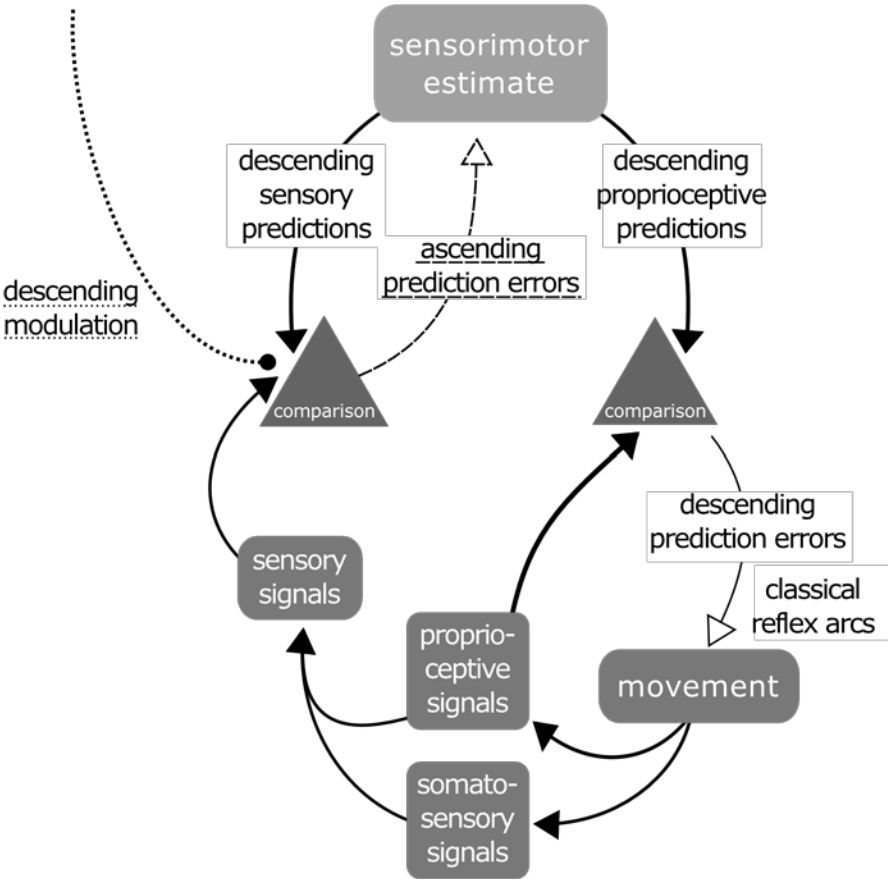

Figure 1: Sensorimotor estimates yield proprioceptive and other sensory predictions. Unless perceptual signals interfere with these estimates (through ascending prediction errors), proprioceptive predictions can be fulfilled through motor reflex arcs. In order for this to be possible, sensory prediction errors have to be attenuated by top-down modulation (for more details, see Brown et al., 2013, p. 418).

movement onset (i.e., they are "turned down", regardless of the signals' actual variance or noise). In the following section, I will argue that this turns peripheral sensory precision estimates into systematically beneficial misrepresentations.

In addition, one could also argue that sensorimotor estimates (according to which, say, the arm is moving) are systematically beneficial misrepresentations prior to movement onset (thanks to Jakob Hohwy for pressing this point in personal communication). A difference to precision estimates is that such estimates are fulfilled through action, so they become true. Therefore, one might be inclined to regard them as representations with a world-to-mind direction of fit (for the notion of a direction of fit see Humberstone, 1992). My desire to have a hot tea right now has the same direction of fit, and its content is counterfactual (because I am not having a tea right now). In fact, the whole point about desires is that they are counterfactual, for I cannot desire something that is actually the case (unless I am unaware that it is the case). Therefore, it would be strange to call a desire misrepresentational. So if sensorimotor estimates prior to movement onset are like desires in this respect, it is at least questionable that they can be called misrepresentational.

Things get more complicated once we consider the question whether representations in PP always have unequivocal directions of fit or not (cf. Wiese, 2014, p. 236). As mentioned above, "active inference induces a circular causality that destroys conventional distinctions between sensory (consequence) and motor (cause) representations." (Friston et al., 2011, p. 138). This suggests that sensorimotor representations prior to movement onset are not like desires (because they do not have an unequivocal world-to-mind direction of fit). Hence, it may be apt to call them misrepresentational. However, a thorough discussion would arguably require a paper on its own, so I will disregard this point in the rest of this paper.

5 Peripheral sensory precision estimates are SBMs

In order to make the case that peripheral sensory precision estimates (PSPEs) that enable action are SBMs, I need to establish four things. On the one hand, it must be shown that they are misrepresentational. On the other hand, it must be shown that they satisfy the three criteria defined in section 2: (1) they are not just acceptable, but useful; (2) they are not just lucky falsehoods (but systematically beneficial); (3) they are not just (tolerable) by-products of other beneficial mechanisms.

5.1 Are PSPEs misrepresentational at the onset of movement?

The challenge entailed by the question in this subsection's heading it not to show that peripheral sensory precision estimates (PSPE) are inaccurate (this should be clear from the above), but that they function as representations in the first place. To show this, their content (i.e., the estimated precision) must be relevant to the role they fulfill according to the computational models by which they are posited (cf. Ramsey, 2007, p. 27).

At first sight, one might suspect that they merely function as switches or causal mediators. During action, they ensure that sensory prediction errors have a low influence on higher-level processing (they "switch sensory signals off"); when no action is performed, they allow sensory prediction errors to have a stronger influence on higher-level processing (they "switch sensory signals on").

However, this is not an adequate description, because sensory prediction error signals are not completely "turned off" during action, but only attenuated. Hence, ascending sensory error signals can still influence processing at higher levels. This is evidenced by the following empirical finding. In the force-matching paradigm (cf. Shergill et al., 2003), where subjects are asked to match a reference force using their own fingers, subjects underestimate the directly self-applied force (it is greater than the matching force they apply in another condition in which they use a joystick to control a robot). The reason for this, according to active inference, is that subjects are less sensitive to sensory input during self-generated movements; and the reason for this is that the estimated precision of PSPEs (related to sensory signals caused by self-generated movements) is decreased.

More specifically, when subjects intend to press their finger on their hand, sensorimotor estimates yield the prediction that a certain pressure is felt on the hand. Prior to motion onset, there will be no such pressure, which results in a prediction error (that has to be attenuated in order to enable movement). When this error is attenuated, proprioceptive predictions can be sustained and will be fulfilled by action (i.e., the finger is pressing on the hand). However, since internal estimates are now less constrained by sensory signals, their predictions become less precise (cf. Brown et al., 2013, p. 422); in particular, predictions regarding how much pressure is actually being applied to the hand become less precise (because the influence sensory prediction errors have on updates is lowered). The result is that subjects are still able to feel the force they are applying with their own fingers, but they have to press stronger in order to be confident that they are pressing hard enough to match the reference force (which in effect means they are applying too much force).

During movement, PSPEs therefore still function as precision estimates (modulating the impact prediction errors have). They do not simply "switch" sensory signals on or off.

5.2 PSPEs satisfy the criteria for SBMs

Note first that modulated PSPEs are not just *acceptable* misrepresentations, but actually beneficial. By enabling action, they fulfill a useful function. Furthermore, they do not enable action by chance, or only under favorable circumstances, but *reliably*. In this sense, they are *systematically* beneficial misrepresentation. Finally, PSPEs are

Penultimate draft. The final publication is available at http://link.springer.com/article/10.1007/s10670-016-9867-x. Read-only version: http://rdcu.be/nZs0.

not just acceptable by-products (of other beneficial mechanisms). By contrast, sensory attenuation (i.e., the phenomenon that subjects are less sensitive to the intensity of sensory signals caused by self-generated movements) is an acceptable by-product of the top-down modulation of PSPEs during action.

Another way to see that modulated PSPEs (which enable action) are not just acceptable by-products is by focusing on the distinction between short-term and long-term benefits of attenuating sensory precision estimates. At first sight, it might seem that PSPEs have only long-term benefits, because they enable action, which helps organisms like us to survive. Note that actions like finding food or shelter are not obviously beneficial in the short run, because they consume energy and can be dangerous. But although short-term effects of action can be harmful, these are usually outweighed in the long run by beneficial effects – so modulated PSPEs might even appear to be merely acceptable by-products of long-term effects. However, attenuating sensory precision estimates is actually more fundamental than this, because it is an *enabling condition* of such long-term benefits (and thus not a by-product). So the reason why modulated PSPEs are beneficial is not that they have immediate or mediate positive effects, but because they are a computational condition of possibility of such effects.[6]

6 Why believe in active inference?

Even if we assume that all of the above has been convincing, we can still question that action is really enabled by SBMs. So far, I have only argued for the conditional claim that action is enabled by SBMs *if* we assume active inference tells a true story about the computational mechanisms underpinning action. Why should we believe that these models are correct?

Note first that active inference coheres well with prominent theories of action representation, viz. common coding (or ideomotor) theories. As shown in section 3, they not only share the assumption that the brain represents actions in terms of their (sensory) consequences, but active inference also *extends* those theories and provides answers to questions left open by ideomotor theories. These questions, as you will recall, were (1) how sensory input is prevented from interfering with motor intentions, and (2) how movements can be imagined or observed without actually being carried out.

Apart from these considerations, there is also empirical support for the story told by active inference. I have already mentioned the force-matching illusion as an example. Active inference models explain why subjects fail to accurately match a reference force using their own fingers, even though they are able to do so when they are applying the force indirectly with a joystick (cf. Shergill et al., 2003).

But active inference can also explain why *schizophrenia patients* show less sensory attenuation (and perform better at the force-matching task, cf. Shergill et al., 2005). Recall that, according to active inference models, sensorimotor estimates must be sustained to initiate action. Unless peripheral sensory prediction errors are attenuated (by decreasing the associated precision estimate through top-down modulation), these error signals will correct sensorimotor estimates. However, all that is necessary, from the point of view of active inference, is that proprioceptive predictions persist in the presence of ascending prediction error. Clearly, there are two ways to change the balance between descending predictions and sensory signals. One way is to attenuate ascending prediction error signals. The other way is to increase the precision estimate associated with sensorimotor estimates. According to active inference, this is what happens in schizophrenia. Brown et al. (2013) have tested this hypothesis in simulations of the force-matching task. The result is that hidden causes are no longer correctly inferred when sensory error signals are not attenuated during action:

> The reason for this false inference or delusion is relatively simple: action is driven by proprioceptive prediction errors that always report less force than that predicted (if they did not, the reflex would not be engaged). However, when sensory precision increases, somatosensory prediction errors become very precise and need to be explained—and can only be explained by falsely inferring an opposing exogenous force. (Brown et al., 2013, p. 423)

Due to the fact that this false inference happens at higher levels, Brown et al. call it a delusion. Furthermore, the simulations replicated the finding that sensory attenuation is diminished in schizophrenia, leading to a more veridical perception of the consequences of self-applied forces (cf. Brown et al., 2013, p. 412). At the same time, delusions are among the positive symptoms of schizophrenia, as the authors note (cf. 2013, pp. 423 f.). Therefore, it is plausible to assume that, in schizophrenia patients, action is not (just) enabled via a descending modulation of sensory precision estimates; instead, precision estimates related to sensorimotor estimates are (also) increased. Active inference can thus account for the performance of neurotypical and of schizophrenia patients in the force-matching task.

Penultimate draft. The final publication is available at http://link.springer.com/article/10.1007/s10670-016-9867-x. Read-only version: http://rdcu.be/nZs0.

Are there better alternatives to the explanation provided by active inference? A prominent competing account is derived from work on motor control that involves *efference copies* and forward models (cf. Wolpert and Flanagan, 2001). A forward model is a system that mimics the input-output behavior of another system (for the difference between forward models and generative models, see Friston, 2011; Pickering and Clark, 2014). For instance, a forward model of an arm receives efferent copies of motor signals and computes how the arm will behave if it receives such signals (or it computes the sensory feedback that will be obtained after the motor signal has been carried out). Forward models can thus be used to compute *predictions* (of how another system will behave or of the feedback it will yield). Crucially, such predictions will be most accurate for self-generated movements (because the motor signal is known for such movements). When external forces influence the target system (say, the arm), there will be a prediction error (because the actual behavior of the arm is not only influenced by the motor signal, but also by some external force that may be unpredictable).

How do accounts that refer to efference copies and forward models explain sensory attenuation and its effects (e.g., the force-matching illusion)? The idea is that the sensory effects of self-generated movements can be predicted well, so there will be a small prediction error. Therefore, the effects will be perceived with less intensity (cf. Blakemore et al., 1999, pp. 554 f.). Brown et al. (2013) criticize this explanation on several grounds:

> Firstly, it is unclear why the intensity of a percept is related to the size of prediction error: prediction errors are used to update predictions, but they do not constitute predictions or percepts per se. [...] Furthermore, the amplitude of prediction error does not seem to be important in determining the level of sensory attenuation: for example, Bäß et al. (2008) show that the predictability of a self-generated sensation does not affect sensory attenuation. [...]
>
> Third, there is a set of results that control theory approach cannot account for. During self-generated movement, sensory attenuation is often noted in response to *externally* generated stimuli (Chapmanet al., 1987; Milne et al., 1988; Rushton et al., 1981; Voss et al., 2008). These stimuli are applied by the experimenter, so they cannot be predicted by the forward model and therefore cannot be attenuated. (Brown et al., 2013, p. 413; citation style adapted)

According to the efference copy explanation, sensory attenuation is a consequence of the *predictability* of sensory signals (i.e., the sensory effects of self-generated movements are highly predictable and therefore yield small prediction errors). According to the active inference explanation, sensory attenuation is a consequence of *top-down estimates* of the precision of sensory signals (i.e., during action, weighted sensory prediction error signals are small because they have very low weights). In order to decide which explanation should be preferred, one should therefore investigate whether the degree of sensory attenuation of a sensory signal correlates positively with the predictability of the signal or not. In the passage quoted above, Brown et al. (2013) already cite some evidence against this correlation. Further evidence against efference copy explanations comes from research on self-tickling.

It is well-known that most people are unable to tickle themselves. Blakemore et al. (1999) provide an efference copy explanation of this phenomenon. Since self-generated movements are highly predictable, self-induced touched is perceived with low intensity. This is why subjects do not have highly intense tickle sensations when they are trying to tickle themselves. Conversely, if self-generated movements are perturbed or delayed (due to external, unpredictable forces), the degree of ticklishness should be significantly higher. This is exactly what Blakemore et al. (1999) found (using a robotic device to control perturbations and delays), but their interpretation of this finding has recently been challenged.

Van Doorn et al. (2015) investigated the alternative interpretation that unpredictable perturbations to self-induced touch only enable self-tickling when those perturbations draw subjects' *attention* to the tickle sensation. So when subjects do not attend to sensory input, they should still be unable to tickle themselves, even though the touch is unpredictable (due to external perturbation).

At first sight, this interpretation might seem completely *ad hoc*. But from the point of view of PP, it makes a lot of sense since PP supports the hypothesis that attention is identical to the process of precision optimization (cf. Feldman and Friston, 2010). According to this view, when we attend to a stimulus, the precision estimate associated with the representation of the stimulus is increased. Hence, *decreasing* peripheral precision estimates at the onset of movement can be described as the process of *attending away* from sensory input. Conversely, when subjects do attend to sensory input, this should lead to less sensory attenuation.

Van Doorn et al. (2015) used a setup that was similar to the one used in Blakemore et al. (1999), but they made sure that subjects were not aware of perturbations to their motion trajectories. Interestingly, they did not find a difference in tickle-ratings in the perturbation condition, i.e., the degree of ticklishness was not judged to be different when there were perturbations to self-generated movements (but for externally generated movements, ratings of ticklishness were stronger, cf. Van Doorn et al., 2015, p. 150). This contrasts with the result reported in Blakemore et al. (1999), where ratings of self-induced tickles were stronger in the presence of perturbations.

As Van Doorn et al. (2015, p. 152) suggest, a crucial difference between the experiments may be that subjects in the earlier study were *aware* of perturbations to their own movements (cf. Blakemore et al., 1999, p. 555). To test this, Van Doorn et al. ran a second experiment in which subjects were told about perturbations and were asked to pay attention to perturbations. As a result, there was no difference in tickle ratings between self- and externally generated movements, irrespective of whether there were perturbations or not (cf. Van Doorn et al., 2015, p. 151). This means that (contrary to what was found in the first experiment), tickle ratings were not stronger for externally generated movements.

Van Doorn et al. (2015) explain this in terms of the amount of attention required to detect perturbations (here, perturbations were more difficult to detect than in the setup used by Blakemore et al.). According to this interpretation, perturbations in the second experiment drew attention away from the sensory input (and from the tickle sensation). Consequently, tickle ratings were lower even in the externally generated condition. Note that this effect poses of puzzle for efference-copy explanations: attention does not make perturbations more predictable, so the tickle ratings should be expected to be higher in the presence of perturbations.

But why do subjects in Blakemore et al.'s study reported stronger tickle sensations when there were perturbations in the self-generated condition? According to the active inference explanation, tickle sensations should be higher when attention is shifted to sensory input, i.e., when the associated precision estimates are higher. At the onset of movement, these precision estimates *must* be low, so it seems impossible to move *and* to attend to sensory input at the same time. Van Doorn et al. suggest the following:

> [T]he introduction of the robot manufacturing the perturbation should itself be modelled internally by the agent as an independently-contributing, modulatory cause of the sensory input; since this intervening cause is external to the agent it may alleviate the need for sensory attenuation of internal signals. (Van Doorn et al., 2015, p. 152)

As I interpret this suggestion, the idea is that internally modelling the robotic device increases the estimated precision associated with sensorimotor estimates. Hence, these estimates can prevail even if (some) attention is shifted to sensory signals (which means that there is less attenuation of the corresponding prediction error). Although this explanation may require further validation, it is a good example of the fruitfulness of active inference, and of predictive processing in general. By contrast, it is doubtful if efference copy explanations can be adapted to the findings presented above.

In sum, these considerations justify the claim that active inference models provide the best explanation of the empirical results considered above. Due to the fact that active inference is part of a rich framework, which also incorporates attention, it yields more specific (and less *ad hoc*) explanations than accounts based on forward models and efference copies. Furthermore, active inference coheres well with prominent ideomotor theories of action, and provides answers to questions left open by such theories. Consequently, there are good reasons to believe that action is enabled by systematically beneficial misrepresentations (SBMs).

7 Replies to objections

Objection 1: Are PSPEs really design features or just acceptable design flaws?

Conceding that peripheral sensory precision estimates (PSPEs) are beneficial (because they enable action), one could suspect that they would be even more beneficial or useful if they were more accurate (thanks to an anonymous referee for suggesting this objection). This would mean that PSPEs at the onset of self-generated movement are not design features, but only acceptable design flaws (or byproducts).

Recall from section 2 that a misrepresentation is an acceptable design flaw when it is (slightly) inaccurate, without significantly impairing the function for which it is used. The functionality of a watch that loses a second a day is not significantly impaired (because we do not usually need more precise specifications). Furthermore, *correcting* the misrepresentation (making it more accurate) would not be disadvantageous (if we disregard possible costs that might go along with making the representation more accurate). It may be tiresome to manually synchronize my slightly imprecise watch with the time indicated by an atomic clock, but if my watch just happened to be more precise, its functionality would not be impaired (on the contrary, it would rather be increased).

Contrast this with possible examples of design features. If having positive illusions about your children turns you into a better parent, then believing that your children have more positive attributes than the average child may be beneficial. What turns this into a design feature is that you become a better parent *regardless* of whether your children actually have the attributes you think they have. Crucially, correcting overly positive misbeliefs about one's children would *diminish* this effect. The point about design-feature misrepresentations is that making them more accurate would impair the beneficial function they have (and not just due to metabolic or other costs that might go along with that).

What about peripheral sensory precision estimates (PSPEs) at the onset of movement? Would they become less useful if they were more accurate? Note that making them more accurate would mean that there would be no (or much less) top-down modulation of PSPEs. All other things being equal, this would result in impairments of movement initiation. If no top-down modulation at all took place, subjects would be virtually unable to generate movements. The reason for this is that sensory prediction errors (signaling that the body is *not* moving) would not be attenuated and would therefore lead to a correction of somatosensory predictions (which, prior to movement onset, must represent that the body *is* moving). So making PSPEs more accurate would impair their function.

This can, in principle, be compensated for by increasing the precision estimates associated with sensorimotor representations. However, as discussed in section 6, this will lead to more severe misrepresentations (delusions), in order to explain away the prediction error: for instance, if internal forces are generated that would normally move my arm and if sensory prediction errors signal that the arm is not moving, this can only be explained by assuming there is an external, counteracting force (see Brown et al., 2013, p. 423). Making PSPEs more accurate would thus either impair their function (i.e., action would be impossible), or it would lead to more severe misrepresentations (delusions). This shows that PSPEs at movement onset are not just acceptable design flaws (of by-products), but design features.

Objection 2: Isn't the modulation of PSPEs just an *attentional* process, according to PP?

Recall that PP suggests we can identify attention with the optimization of precision estimates. So when a precision estimate is increased or decreased, this means the subject is attending to or away from something. Prior to movement onset, decreasing PSPEs just means that subjects shift their attention away from sensory input. In fact, if attention is the process of optimizing precision estimates, then PSPEs must be optimal in some sense, so why should they be called misrepresentational (thanks to Andy Clark for suggesting this objection)?

There are at least two ways in which this objection can be interpreted. On the one hand, it can be read as another way of describing the sense in which PSPEs are design features: they are the result of an *optimization process*, and still they are inaccurate. They are optimal because they enable action (and avoid more severe misrepresentations), and they are misrepresentational because they do not reflect the actual precision of sensory signals.

On the other hand, the objection could be interpreted as claiming that the primary function of modulating precision estimates is to allocate attention. When a precision estimate is increased, its main function is not to indicate that the noise level (or uncertainty) associated with some signal is decreasing, but to draw attention to a certain stimulus or region. Attention, however, is not something which could be described as a misrepresentation. Hence, PSPEs should not be described as misrepresentations.

This version of the objection conflates the process of attention with representations (precision estimates) that are postulated as part of the computational mechanisms underpinning attention. Attention itself is not a representation, but computational models of what happens when attention is allocated posit representations. Furthermore, modulating precision estimates does not serve the function of allocating attention; it is *identical to* the process of allocating attention, under the computational descriptions provided by PP. But these descriptions involve representations (precision estimates), and some of them misrepresent systematically.[7]

8 Conclusion

I have argued that low-level sensory precision estimates, as posited by active inference (in predictive processing), are systematically misrepresentational. In spite of this, they are beneficial for the agent, because they enable action. This means they are not just tolerable by-products or acceptable design flaws, but design features (in the sense of McKay and Dennett, 2009). Furthermore, I have argued that there are good reasons to believe that active inference provides an accurate account of how action is enabled. First of all, it is compatible with prominent ideomotor theories of action representation. Secondly, it extends those theories by specifying the computational mechanisms that enable (or inhibit) action. Thirdly, it can explain empirical results related to sensory attenuation. Fourthly, these explanations are (as I have tried to show) more convincing than competing efference copy explanations.

This suggests not only that peripheral sensory precisions estimates are systematically misrepresentational. Since they enable action, and since this is probably among the most important functions fulfilled by neural structures, systematically beneficial misrepresentations may lie at the heart of our neural architecture.

Acknowledgements

I am highly grateful to Jakob Hohwy for providing very detailed and helpful feedback on an earlier version of this paper. Some of the ideas present in this paper were presented in a talk I gave at a workshop in the context of the Carnap lectures in March 2014 in Bochum. I am grateful to the audience of that workshop, special thanks to the organizers Albert Newen and Tobias Schlicht, and to Peter Brössel and Daniel Dennett. A previous version of this paper was presented at the Journal Club of the theoretical philosophy group at Mainz University, organized by Thomas Metzinger; thanks to all participants. Another version was presented at Spindel 2014 in Memphis. Thanks to the organizer Shaun Gallagher, as well as to the audience of that conference. Thanks to two anonymous referees for providing a number of very helpful comments. Part of the work on this paper was supported by a scholarship of the Barbara Wengeler foundation.

References

Adams, R. A., Shipp, S. & Friston, K. J. (2013). Predictions not commands: active inference in the motor system. *Brain Structure and Function*, *218*(3), 611–643. DOI 10.1007/s00429-012-0475-5

Bäß, P., Jacobsen, T. & Schröger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology*, *70*(2), 137–143.

Blakemore, S.-J., Frith, C. D. & Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience*, *11*(5), 551-559. DOI 10.1162/089892999563607

Botvinick, M. & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature*, *391*(6669), 756–756.

Brown, H., Adams, R., Parees, I., Edwards, M. & Friston, K. (2013). Active inference, sensory attenuation and illusions. *Cognitive Processing*, *14*(4), 411-427. DOI 10.1007/s10339-013-0571-3

Chapman, C., Bushnell, M., Miron, D., Duncan, G. & Lund, J. (1987). Sensory perception during movement in man. *Experimental Brain Research*, *68*(3), 516–524.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. DOI 10.1017/S0140525X12000477

Clark, A. (2015).Radical predictive processing. *The Southern Journal of Philosophy*, *53*, 3-27. DOI 10.1111/sjp.12120

Clark, A. (2016). *Surfing uncertainty: Prediction, action, and the embodied mind.* (New York: Oxford University Press.)

Feldman, H. & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, *4*. DOI 10.3389/fnhum.2010.00215

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. DOI 10.1038/nrn2787

Friston, K. (2011). What is optimal about motor control? *Neuron*, *72*(3), 488–498. DOI 10.1016/j.neuron.2011.10.018

Friston, K., Adams, R., Perrinet, L. & Breakspear, M. (2012a). Perceptions as hypotheses: saccades as experiments. *Frontiers in Psychology*, *3*(151). DOI 10.3389/fpsyg.2012.00151

Friston, K., Mattout, J. & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, *104*(1-2), 137–160. DOI 10.1007/s00422-011-0424-z

Friston, K., Samothrakis, S., & Montague, R. (2012b). Active inference and agency: optimal control without cost functions. *Biological Cybernetics, 106*(8), 523–541. DOI 10.1007/s00422-012-0512-8

Friston, K. J., & Stephan, K. E. (2007). Free-energy and the brain. *Synthese, 159*(3), 417-458. DOI 10.1007/s11229-007-9237-y

Gigerenzer, G., Todd, P. M. & the ABC Research Group. (1999). *Simple heuristics that make us smart*. (New York (a.o.): Oxford University Press.)

Herbart, J. F. (1825). *Psychologie als Wissenschaft neu gegründet auf Erfahrung, Metaphysik und Mathematik. Zweiter, analytischer Teil.* (Königsberg: Unzer.)

Hohwy, J. (2013). *The predictive mind*. (Oxford: Oxford University Press.)

Hommel, B. (2015). The theory of event coding (TEC) as embodied-cognition framework. *Frontiers in Psychology*, *6*. DOI 10.3389/fpsyg.2015.01318

Hommel, B., Müsseler, J., Aschersleben, G. & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*, 849–878. DOI 10.1017/S0140525X01000103

Humberstone, I. L. (1992). Direction of fit. *Mind*, *101*(401), 59–83.

Hurley, S. L. (1998). Vehicles, contents, conceptual structure, and externalism. *Analysis*, *58*(1), 1–6.

James, W. (1890). *The principles of psychology*. (New York: Henry Holt.)

Lotze, R. H. (1852). *Medicinische Psychologie oder Physiologie der Seele*. (Leipzig: Weidmann'sche Buchhandlung.)

McKay, R. T. & Dennett, D. C. (2009). The evolution of misbelief. *Behavioral and Brain Sciences*, *32*(6), 493–510. DOI 10.1017/S0140525X09990975

Mendelovici, A. (2013). Reliable misrepresentation and tracking theories of mental representation. *Philosophical Studies*, *165*(2), 421–443. DOI 10.1007/s11098-012-9966-8

Milne, R., Aniss, A., Kay, N. & Gandevia, S. (1988). Reduction in perceived intensity of cutaneous stimuli during movement: a quantitative study. *Experimental Brain Research*, *70*(3), 569–576.

Pickering, M. J. & Clark, A. (2014). Getting ahead: forward models and their place in cognitive architecture. *Trends in Cognitive Sciences*, *18*(9), 451–456. DOI 10.1016/j.tics.2014.05.006

Prinz, W. (1990). A common coding approach to perception and action. (In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action* (pp. 167–201). Berlin; Heidelberg: Springer.)

Ramsey, W. (2007). *Representation reconsidered*. (Cambridge: Cambridge University Press.)

Rushton, D., Roghwell, J. & Craggs, M. (1981). Gating of somatosensory evoked potentials during different kinds of movement in man. *Brain*, *104*(3), 465–491.

Sharot, T. (2011). The optimism bias. Current Biology, 21(23), R941–R945. DOI 10.1016/j.cub.2011.10.030

Shergill, S. S., Bays, P. M., Frith, C. D. & Wolpert, D. M. (2003). Two eyes for an eye: The neuroscience of force escalation. *Science*, *301*(5630), 187. DOI 10.1126/science.1085327

Shergill, S. S., Samson, G., Bays, P. M., Frith, C. D. & Wolpert, D. M. (2005). Evidence for sensory prediction deficits in schizophrenia. *American Journal of Psychiatry*, *162*(12), 2384–2386.

Shipp, S., Adams, R. A. & Friston, K. J. (2013). Reflections on agranular architecture: predictive coding in the motor cortex. *Trends in Neurosciences*, *36*(12), 706–716. DOI 10.1016/j.tins.2013.09.004

Taylor, S. E. (1989). *Positive illusions: Creative self-deception and the healthy mind*. New York, NY: Basic Books.

Trivers, R. (2000). The elements of a scientific theory of self-deception. *Annals of the New York Academy of Sciences*, *907*, 114–131. DOI 10.1111/j.1749-6632.2000.tb06619.x

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124-1131. DOI 10.1126/science.185.4157.1124

Van Doorn, G., Paton, B., Howell, J. & Hohwy, J. (2015). Attenuated self-tickle sensation even under trajectory perturbation. *Consciousness and Cognition*, *36*, 147–153. DOI 10.1016/j.concog.2015.06.016

Voss, M., Ingram, J. N., Wolpert, D. M. & Haggard, P. (2008). Mere expectation to move causes attenuation of sensory signals. *PLoS ONE*, *3*(8), e2866. DOI 10.1371/journal.pone.0002866

Weinstein, N. D. (1980). Unrealistic optimism about future life events. Journal of Personality and Social Psychology, 39(5), 806–820. DOI 10.1037/0022-3514.39.5.806x

Wenger, A. & Fowers, B. J. (2008). Positive illusions in parenting: Every child is above average. *Journal of Applied Social Psychology*, *38*(3), 611–634. DOI 10.1111/j.1559-1816.2007.00319.x

Wiese, W. (2014). Jakob Hohwy: The predictive mind. *Minds and Machines*, *24*(2), 233–237. DOI 10.1007/s11023-014-9338-6

Wolpert, D. M. & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*(18), R729–R732.

Penultimate draft. The final publication is available at http://link.springer.com/article/10.1007/s10670-016-9867-x. Read-only version: http://rdcu.be/nZs0.

Zehetleitner, M. & Schönbrodt, D. F. (2015). When misrepresentation is successful. (In T. Breyer (Ed.), *Epistemological dimensions of evolutionary psychology* (pp. 197–221). New York, NY: Springer New York.) DOI 10.1007/978-1-4939-1387-9_10

---

[1] Positive illusions are false beliefs that depict states of affairs as overly positive (in some cases, this may systematically promote well-being and mental health, cf. Taylor, 1989). A particular example is the *optimism bias* (also called *unrealistic optimism*, cf. Weinstein, 1980), which refers to overly positive expectations about future events. As such, an optimism bias need not be beneficial for the agent, but at least a moderate optimism bias seems to be (cf. Sharot, 2011, p. R944).

[2] The earliest pre-cursor of ideomotor theories seems to be the view put forward in Herbart (1825, pp. 464 f.).

[3] Technically, both active inference and predictive processing are corollaries of Karl Friston's free-energy principle (cf. Friston, 2010). The free-energy principle provides a general formulation of self-organization where, in the context of the brain, prediction errors can be regarded as a proxy for free energy.

[4] Perceptual and active inference are sometimes described as distinct, but complementary processes. However, "active inference" is also used as a generic term for the computational processes underpinning both action and perception (cf. Friston et al., 2012b, p. 539). This is the primary sense in which the term is used in this paper.

[5] Note here a distinction between ascending and descending prediction errors. The ascending prediction errors from the spinal cord are those that revise estimates in the central nervous system, while descending prediction errors are sent to the periphery to activate classical motor reflexes. Attenuation of ascending prediction errors enables movement, while attenuation of both ascending and descending prediction errors is necessary to prohibit echopraxia, mimicry, or mirror movements during action observation. Thanks to an anonymous referee for pointing this out.

[6] From the point of view of Karl Friston's free-energy principle, the relevance of attenuating sensory precision is even more obvious. Active inference is here regarded as a means to minimize free energy by changing sensory samples. Under quite general assumptions, minimizing free energy is entailed by minimizing prediction error. So the reduction of prediction errors in active inference necessarily depends upon resampling the world and, by

implication, the attenuation of sensory precision. This means, in the long-term, the misrepresentation (attenuation) of sensory precision reduces prediction errors and is quintessentially beneficial for any sentient agent, because it enables active inference and thereby enables the agent to stay within viable regions of its state space. This is the premise of the free-energy principle, in which prediction errors can be regarded as free energy – and both are proxies for negative (Bayesian) model evidence. I am grateful to an anonymous referee for emphasizing this point.

[7] As an anonymous reviewer remarked, if one identifies attention with predictions of precisions (not with the process of optimizing these predictions), and if the precision predicted does not correspond to the actual signal-to-noise ratio of sensory inputs (as in sensory attenuation or spatial attention) then, by definition, attention can be called a beneficial misrepresentation.