Reply

Weighting models and weighting factors

Gottfried Vosgerau, Matthis Synofzik

Department of Philosophy, Heinrich-Heine-Universität, Düsseldorf, Germany
Department of Neurodegenerative Diseases, Hertie-Institute for Clinical Brain Research, University of Tübingen, Germany
German Research Center for Neurodegenerative Diseases, Tübingen, Germany

A R T I C L E   I N F O

Article history:
Received 20 September 2011
Available online 24 October 2011

Keywords:
Agency
Comparator model
Schizophrenia
Optimal cue integration
Bayesian cue integration
Multifactorial weighting model

A B S T R A C T

We defend our multifactorial weighting model of the sense of agency and our critique of the comparator model (Synofzik, Vosgerau, & Newen, 2008) against the critiques that have been brought forward by Carruthers (2012) and Wong (2012). Building on the specification of our model that emerges from this response, we will suggest a distinct mechanism how weighting of different agency factors might work: internal and external agency cues are constantly weighted according to their reliability in a given situation. Thus, the weighting process underlying the sense of agency might follow the principles of optimal cue integration. We review recent empirical evidence for this hypothesis, demonstrating that the multifactorial weighting model is not only testable, but has in fact already received first empirical support.

© 2011 Elsevier Inc. All rights reserved.

1. Weighting different comparators and different steps

In his remarks, Wong (2012) argues that the multifactorial weighting model (MWM) leaves out one important aspect of agency, namely the sense of controlling an action. Accordingly, he criticizes our (Synofzik, Vosgerau, & Newen, 2008) arguments against the comparator model (CM) since it leaves out comparator number 2 (see Fig. 1 in Wong, 2012) which is responsible for control. It is true that we concentrated on comparator number 3 in our paper. However, as stated in the paper, we did so for specific reasons: “In the following critique we will exemplarily focus on the latter comparator (comparator 3) since it is this comparator that is (i) responsible for attributing self-agency to sensory events and that is (ii) currently held responsible for disruptions of agency in schizophrenia patients (Frith, 2005; Lindner, Thier, Kircher, Haarmeier, & Leube, 2005; Shergill, Samson, Bays, Frith, & Wolpert, 2005). In contrast, the former comparator (comparator 2) (i) cannot explain self-agency for sensory events (since no kind of sensory signal is entered in the comparison process), (ii) is more speculative in character and less well empirically established and (iii) is not primarily used anymore to explain delusions of control in schizophrenia (Frith, 2005). However, most of the following critique regarding comparator 3 applies to comparator 2 as well.” (Synofzik et al., 2008, p. 221). We would like to add that even most of the current advocates of the CM relate the sense of agency to comparator 3 and not to comparator 2 (see e.g. Carruthers, 2012, figure).
Apart from these points, one of the advantages of the MWM is exactly that it regards all of the factors entering comparator 2 (namely internal predictions and motor goals, including “control”) and comparator 3 (internal predictions and sensory feedback) as contributing factors to the sense of agency. In this respect—following the argumentation of Wong—the MWM seems even more adequate than the CM to account for the sense of agency: it is not restricted to only a particular subset of cues (like e.g. the comparator 2 or the comparator 3, respectively) and does not need to postulate the existence of several comparators (evoking the problem of how to combine the different comparator outputs). Instead, it flexibly integrates a large variety of cues, depending on their availability and reliability. Thus, the notion of a MWM is well consistent with the findings of an illusionary movement in subjects receiving direct electrical stimulation of posterior parietal cortex with high intensities (Desmurget et al., 2009). Here, premotor signals (like motor intentions and internal predictions) suffice to install a sense of agency even in the absence of sensory feedback, since they appear to be strong (and thus reliable) enough.

The second point of Wong concerns a general worry about the two-step account of the MWM differentiating between the feeling of agency and the judgment of agency. Although there are some misconceptions on Wong's side on how we define the sense of agency (neither does the feeling of agency need to be conscious nor does the sense of agency necessarily involve an explicit judgment of agency, especially in unambiguous circumstances; cf. Synofzik et al., 2008, pp. 227–229), he generally acknowledges the distinction to be an important and useful one. However, he is not convinced that all cases of pathological agency attributions to oneself necessarily involve both steps. His suggestion is—following the Bayesian account proposed by Fletcher and Frith (2009)—that judgments can arise directly from perception. This suggestion, however, does not contradict our two-step model. In fact, we explicitly state that “the feeling of agency might be so strong that the belief formation mechanism does not need to be further instantiated” (Synofzik et al., 2008, p. 229), meaning that in these cases the judgment is exclusively based on the feeling. Similarly, perceptual agency cues might be so strong and irresistible that they are the only basis for a strong feeling of agency, thereby directly triggering and constraining the according agency judgment. Therefore, the second step in our model is not necessarily a step of “active reasoning” (as Wong seems to construe it). Nevertheless, insofar a belief is formed at all (however direct this formation is), there is a transition from a non-conceptual feeling to a conceptual belief (a second step).

Moreover, even Fletcher and Frith (2009) themselves allow for a certain kind of two-step (or even multi-step) model: First, predictions about the effects of movements are distorted, leading to abnormal experiences in the patients. Second, the belief system of the patient is adjusted to accommodate for the abnormal experiences (p. 55f). Accordingly, Chris Frith himself concludes “that the two-factor account of delusions would be consistent with a Bayesian account” (Frith, 2012, p. 56–58).

One further remark on the prediction model of Fletcher and Frith: the predictions which are postulated by Fletcher and Frith are not predictions about agency; they are predictions about the effects of movements. The fact that they can be viewed as beliefs (namely as beliefs about future states of the world) does not constitute a case against a two-step model of agency, since the judgment of agency is not a belief about future states of the world (a prediction) but is a belief about the causal “source” of these changes. Thus, it cannot be construed as being a prediction of anything—if, at all, it is a “postdiction” about the cause of certain events.

Moreover, Wong agrees that a two-step model is apt in most cases. If it is also able to explain odd cases (which it does), this is a good reason to accept the model even for cases in which a one-step account would be possible. The reason is that it is more parsimonious to have one model that explains all cases than to have two models which apply to different cases. Nevertheless, it is important not to “overintellectualize” the second step: Although this step is one of judgment formation, it does not need to be conscious or to involve complex inferences—as in most everyday cases, this step will be automatic, quick, effortless, and not based on the conscious consideration of different factors. Indeed, the same holds for pathological cases.

### 2. Weighting imprecise models

Comparing the Frithian CM of the sense of agency with the MWM, Carruthers (2012) argues that the CM can be adopted to explain every case of unusual agency experiences so far reported in the literature. Or in other words, the CM has never empirically proven to be wrong. However, he admits that for many cases the CM has to be adjusted, and he proposes various—partly even contrary—adjustments for different cases. Some of the proposed adjustments seem to be rather speculative in nature, being only barely supported by direct empirical evidence. For example, with respect to phantom limbs, the author speculates that these subjects (i) form false perceptual representation of one's limbs, (ii) which can act as a continuing learning signal (iii) in order to establish precise internal predictions (Carruthers, 2012). Other adjustments are very indirect and lose the elegant parsimony which made the original CM so attractive. For example, with respect to the adjustment of the CM to priming studies, the authors himself explicitly acknowledges that “the amount of modification to the [comparator] model needed is becoming incredibly large and none of these modifications is predicted by the initial [comparator] model.” (Carruthers, 2012, p. 32–47). Correspondingly, it remains questionable whether it is indeed possible to integrate all different adjustments into a coherently adjusted CM. This question should not be discussed here. The important point is that the CM is not precise enough to be a successful model: in itself, it does not specify a number of problems, thus making various different adjustments possible and necessary.

The MWM is also able to explain all of the cases cited by Carruthers—and probably in a much more direct and parsimonious way. However, according to him, the MWM has one disadvantage, namely that it is even more imprecise than the CM:
No case can be conceived of for which the MWM would not have an explanation ready. Or in other words, it is not falsifiable. Therefore, Carruthers concludes, the CM is still the better model.

In order to evaluate this conclusion, we need to take a closer look at the ways in which both models are imprecise: The CM does not specify, for example, which type of sensory feedback is compared with the prediction. Apparently, no kind of sensory feedback can be shown to be necessary or sufficient to elicit the sense of agency (Synofzik et al., 2008). Thus, the CM fails to specify the factors with which it works. In order to overcome this failure, we would need another model to explain what factors are selected in which context. Exactly this is done by the MWM (Synofzik et al., 2008), which specifies different agency cues on different levels that are weighted according to their relevance in the specific context. (Note that the idea of comparing internal predictions with sensory feedback is incorporated into the MWM as only one factor among others.)

Does it really follow that the MWM is not testable? The key to answer this question is the notion of factor. A factor is something stable, something that produces predictable stable effects if modulated. This is exactly not the case for the “factors” advocated for by the CM: A change in visual feedback, for example, does not produce predictable effects in the sense of agency as there are cases in which it has effects and others in which it does not have effects (Synofzik et al., 2008, pp. 223–224). One way to get around this problem is to assume that factors are weighted depending on the context. Such a weighting has to be stable as well in order to count as a factor itself (a meta-factor, if you wish). If it is not stable across different but comparable cases, it is indeed useless for a scientific model since it would be unpredictable (and hence not testable). In other words: The MWM is confined to the claim that in comparable situations the weighting of a specific factor in one subject (e.g., visual feedback) is always the same. If it could be shown that this is not the case, i.e. that we have to assume different weightings in one and the same experimental condition (other factors equal), the MWM would be falsified, since in this case the weights of the different agency cues could not be understood as factors themselves.

Moreover, further empirical plausibility for the MWM comes from recent experimental data: The study by Wilke and colleagues (2012) shows that a comparison between internal predictions and sensory feedback does not suffice to explain the perception of one’s own actions: in addition to these two cues, the perception of one’s actions is also modulated by external cues presented post hoc, here: the affective valence of action outcomes. This experiment could have provided evidence against the MWM: if the perception of one’s own actions was not modulated by additional cues, but fully determined by the comparison between the predicted and the actual sensory consequences, the assumption of a MWM would be much less plausible. Similarly, a study by Moore, Wegner, and Haggard (2009) showed that, if consistent primes were available, even effects of passive movements are adopted to one’s own agency. Thus, internal predictions (which are only issued in case of active movements) are not even necessary to induce a sense of agency, but external cues (here: primes) can substitute it. Again, this experiment could have provided evidence against the MWM: if the self-registration of action effects was not inducible by external cues, but only by comparison between the predicted and the actual sensory consequences, the assumption of a MWM would be much less plausible.

3. Bayesian optimal cue integration

However, we agree with Carruthers that the question of how weights are assigned to different cues was not tackled in sufficient detail in our initial paper (Synofzik et al., 2008). A proposal has emerged since then: each agency cue is weighted according to its relative reliability in a given situation (Synofzik, Vosgerau, & Lindner, 2009; Synofzik & Voss, 2010; Synofzik et al., 2010). The reliability of a cue would be low if its variance is high; in turn, its reliability would be high if it is present in a very salient way and/or highly precise. This notion follows the framework of optimal cue integration, according to which no single information signal is powerful enough to convey an adequate representation of a certain perceptual entity under all everyday conditions. Instead, depending on the availability and reliability of a certain information cue, different combination and integration strategies should be used to frame the weighting of sensory and motor signals. Thus, optimal cue integration might not only allow robust perception of the world (Ernst & Banks, 2002; Ernst & Bülthoff, 2004) and efficient sensorimotor learning (Kording & Wolpert, 2004), it could also provide the basis for subjects’ robust, and at the same time flexible, agency experience in variable contexts.

Hence, the MWM becomes testable also with respect to the weighting mechanism: the weighting of each cue should be shifted according to its reliability. For example, if internal predictions are missing, external cues (like e.g. primes) should receive a higher weight for determining one’s sense of agency. This idea has received preliminary indirect empirical support from the aforementioned study by Moore et al. (2009). More specifically, the variance within one agency cue should be directly related to the over-reliance on another. This fact has been exemplarily demonstrated in schizophrenia patients: the more imprecise their internal predictions, the stronger their weighting of visual feedback signals (Synofzik et al., 2010).

Acknowledgment

G.V. and M.S. were supported by a grant of the Volkswagen Stiftung “Grounding thoughts in actions (ThinkAct)” (Az. VW II/85 068,158).

\(^{2}\) In case the cue is absent, its value is 0, such that the product of cue value and weight will be 0 whatever the weight is. Therefore, in such cases the value of the weight is irrelevant and does not need to be specified.
References


