EBERHARD KARLS UNIVERSITÄT TÜBINGEN

It was (not) me: Causal Inference of Agency in goal-directed actions

TF Beck[1,2,3], C Wilke[1], B Wirxel[1,3], D Endres*[1,2], A Lindner[1,3], MA Giese*[1,2,3]

*) Equal contribution. 1) University Clinic Tübingen, Dept. of Cognitive Neurology, CIN Tübingen, Hertie Institute for Clinical Brain Research. 2) Section for Theoretical Sensomotorics. 3) BCCN Tübingen. Germany.

Contact: tobias.beck@uni-tuebingen.de, www.compsens.uni-tuebingen.de, www.bccn-tuebingen.de

Hertie Institute for Clinical Brain Research
CIN Werner Reichardt Centre for Integrative Neuroscience
bccn tübingen

Introduction

The perception of own actions is affected by both visual information and predictions derived from internal forward models [1]. The integration of these information sources depends critically on whether visual consequences are associated with one's own action (sense of agency) or with changes in the external world unrelated to the action [2, 3] and the accuracy of integrated signals [4, 5]. The attribution of percepts to consequences of own actions should thus depend on the consistency between internally predicted and actual visual signals.

Goal of this work is to develop quantitative theories for the influence of the sense of agency on the fusion of perceptual signals and predictions derived from internal forward models. Our work exploits *graphical models* as central theoretical framework.

Experiment

In order to study the attribution of sensations to consequences of own actions, we investigated the effect of the consistency between internally predicted and actual sensory consequences using a virtual reality setup. Participants were seated in front of a horizontal board on which their right hand was placed with the tip of the index finger positioned on a haptic marker, representing the starting point for each trial. Participants were instructed to execute straight, fast (quasi-ballistic) pointing movements of fixed amplitude to a circle, which was briefly flashed before the execution of the pointing movements. The target points on the circle could be chosen freely by the participant within the upper right quadrant of the circle. No explicit visual target was presented, and the hand was obstructed from the view of the participants. Visual feedback about the direction for the peripheral part of the movement was provided by a cursor that moved radially towards the circle. The movement of the cursor was computed online from the hand movements of the participants, adding rotations within the plane of the board.

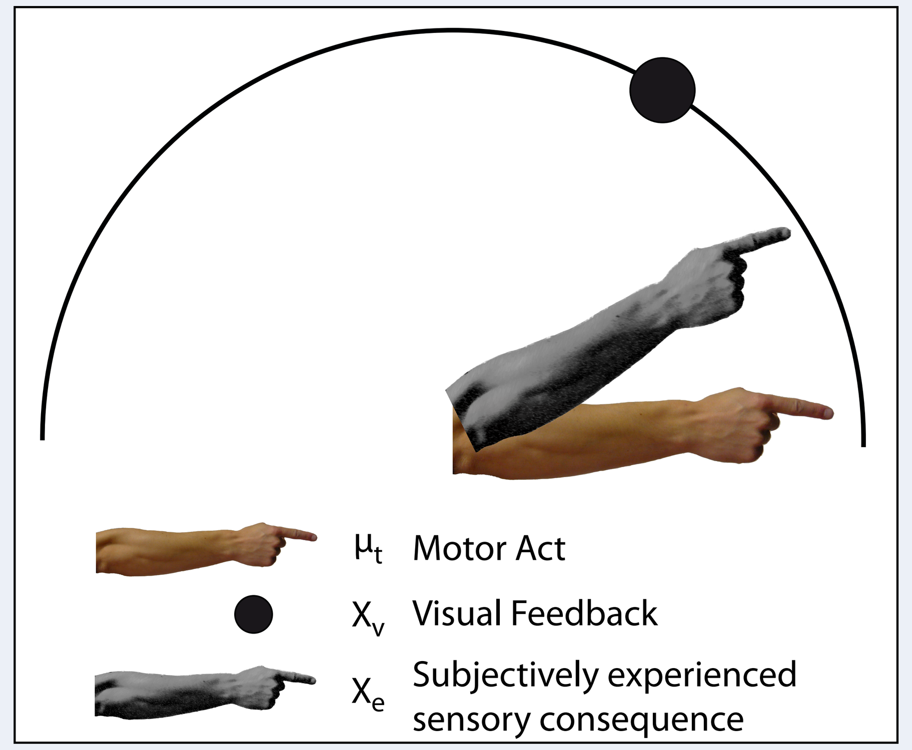

Fig. 1) Experimental paradigm and relevant variables.

The true movement of the fingertip in the plane was captured by an ultrasound tracking system. Visual feedback (cursor movement) was either veridical, or rotated against the true direction of the hand movement by predefined angles (i.e. $\pm 5°$, $\pm 10°$, $\pm 20°$, $\pm 40°$). These offset angles were randomized over the whole experiment in order to minimize the effects of trial-by-trial adaptation. After each trial participants were asked to report the subjectively experienced direction of the executed hand movement X_e by placing a mouse-cursor into that direction. Figure 1) shows an example: The true motor act is pointing in a direction that deviates from the subjectively experienced motion direction (indicated in gray). The visual feedback (indicated in black) deviates from the true motion direction by a predefined offset (order randomized over trials).

Graphical Model

We compared two probabilistic models: Both include a binary random gating variable (*agency*) that models the sense of 'agency'; that is the belief that the visual feedback is influenced by the subject's own motor action. The first model, (Fig. 2), assumes that both the visual feedback X_v and the internal motor state estimate X_e are directly caused by the (unobserved) intended motor state X_t, where μ_t is the true motor action as measured by the tracking system. The second model assumes instead that the expected visual feedback X_v depends on the perceived direction of the own motor action X_e (Fig. 3).

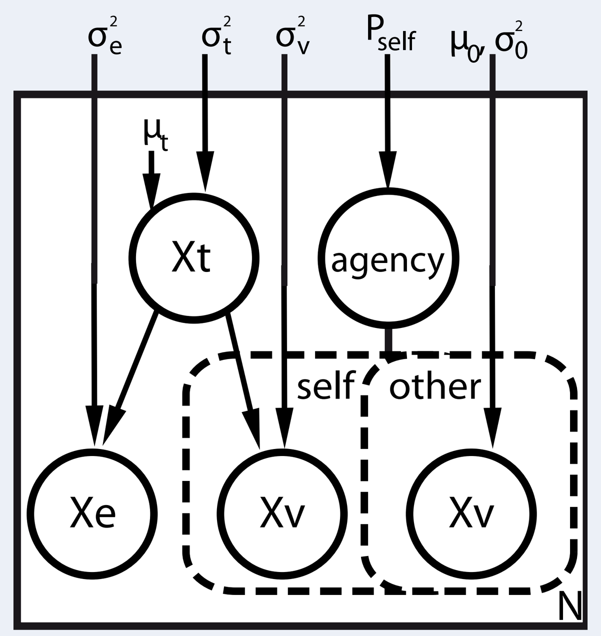

Fig. 2) Model 1.

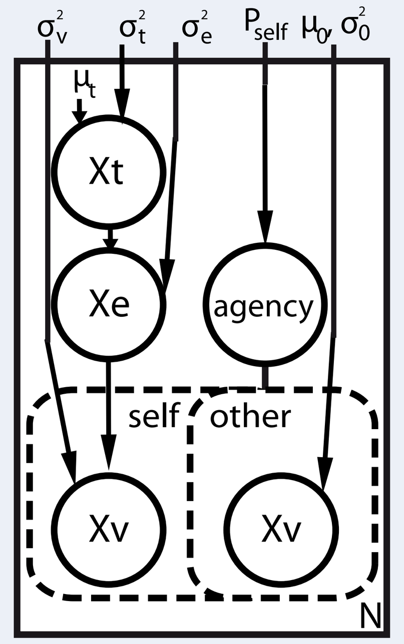

Fig. 3) Model 2.

The variables representing internal forward models and proprioceptive feedback are not shown in the graphical models, as no quantifiable statements about them can be derived from our experimental paradigm. Therefore, these variables are treated as unobserved and their effects are subsumed by the statistics of the internal motor state variable X_e. From the subject's perspective the only observed variables are μ_t, via proprioception, and X_v. The parameters that need to be estimated are the prior statistics for each variable, e.g. for model 1:

$$X_t \sim \mathcal{N}\left(X_t|\mu_t, \sigma_t^2\right);\ X_e \sim \mathcal{N}\left(X_e|X_t, \sigma_e^2\right);\ X_v|agency{=}self \sim \mathcal{N}\left(X_v|X_t, \sigma_v^2\right) \quad (1)$$

$$X_v|agency{=}other \sim \mathcal{N}\left(X_v|\mu_0, \sigma_0^2\right);\ agency \sim \text{Bernoulli}\left(agency|P_{self}\right) \quad (2)$$

In the graphical model, all parameters but μ_t are located outside the plate, signifying their constancy over trials, while the variables inside the plate change for each trial.

Parameter Estimation (1)

Maximum Likelihood

In order to learn the parameters $\Theta = \{P_{self}, \sigma_e^2, \sigma_v^2, \sigma_t^2, \sigma_0^2, \mu_0\}$ it is necessary to maximize the likelihood of the (independent) measurements $X_e^{(i)}$ over all N trials,

$$\arg\max_{\Theta} \prod_i^N \mathrm{p}\left(X_e^{(i)}|X_v^{(i)}, \mu_t^{(i)}, \Theta\right). \quad (3)$$

For improved computational performance, maximizing the log-likelihood is common practice. Optimization was based on an active set method (MATLAB Optimization Toolbox). To avoid local minima, the algorithm was started with 10,000 randomly sampled parameter configurations, selecting the best optimum over this sample. The marginal probability in (3) is a sum over *agency*, conditioned on $\{X_v^{(i)}, \mu_t^{(i)}, \Theta\}$. For sake of readability, the conditioning on $\{\mu_t^{(i)}, \Theta\}$ and the superscript (i) are omitted.

$$\mathrm{p}\left(X_e|X_v\right) = \sum_{agency} \mathrm{p}\left(X_e, agency|X_v\right) \quad (4)$$

The summands in (4) can be calculated by marginalizing X_t out of the joint probability and conditioning on X_v by dividing by the marginal probability $\mathrm{p}(X_v)$, giving

$$\mathrm{p}\left(X_e, agency|X_v\right) = \frac{\int \mathrm{d}\,X_t\, \mathrm{p}\left(X_t, X_e, X_v, agency\right)}{\sum_{agency} \int \mathrm{d}\,X_e \int \mathrm{d}\,X_t\, \mathrm{p}\left(X_t, X_e, X_v, agency\right)}. \quad (5)$$

Parameter Estimation (2)

Only equations for model 1 are shown here, as the derivations for model 2 are similar. The joint probability of (5) can be calculated by exploiting the factorization described by the graphical model, giving

$$\mathrm{p}\left(X_t^{(i)}, X_e^{(i)}, X_v^{(i)}, agency^{(i)}|\mu_t^{(i)}, \Theta\right) = \mathrm{p}\left(X_t^{(i)}\right) \mathrm{p}\left(X_e^{(i)}|X_t^{(i)}\right) \mathrm{p}\left(X_v^{(i)}|X_t^{(i)}, agency^{(i)}\right) \mathrm{p}\left(agency^{(i)}\right). \quad (6)$$

With the probability distributions in eqs. (1) and (2) and eqs. (3) through (6), the parameters were estimated using the above described maximum likelihood optimization criteria. Results of representative subjects are given below.

Bayesian Learning

Instead of just computing point estimates via maximum likelihood, estimation of the complete parameter posterior distribution will allow us to determine whether the parameters are well-constrained by the data, or whether more data need to be recorded. Parameter estimation using Bayesian learning cannot be solved analytically anymore. One approximation approach is the optimization of the variational bound (Gibbs free energy, [6]), $\mathcal{L}$, commonly done by EM steps to alternately optimize for the variational posterior $\mathrm{q}(agency)$ and the posterior hyperparameters $(\alpha, \beta, \kappa, \lambda)$. Within the exponential family conjugate prior framework in [6], priors conjugate to the parameter probability distributions can be defined:

$$P_{self} \sim \text{Beta}\left(P_{self}|a_s, b_s\right), \quad \sigma_t^2 \sim \Gamma^{-1}\left(\sigma_t^2|a_t, b_t\right), \quad \sigma_e^2 \sim \Gamma^{-1}\left(\sigma_e^2|a_e, b_e\right), \quad \sigma_v^2 \sim \Gamma^{-1}\left(\sigma_v^2|a_v, b_v\right), \quad \mu_0, \sigma_0^2 \sim \mathcal{N}\text{-}\Gamma^{-1}\left(\mu_0, \sigma_0^2|l_0, k_0, a_0, b_0\right) \quad (7)$$

The lower bound $\mathcal{L}$ for variational inference can be calculated with (7) from (6) using Jensen's inequality and the Kullback-Leibler divergence, giving

$$\mathcal{L} = \int \mathrm{d}\,X_t^{(1)} \ldots \int \mathrm{d}\,X_t^{(N)} \sum_{agency^{(1)}} \cdots \sum_{agency^{(N)}} \prod_{i=1}^{N} \left[\mathrm{q}\left(agency^{(i)}\right) \mathrm{q}\left(X_t^{(i)}\right) \mathrm{q}\left(\Theta\right)\right] \sum_{i=1}^{N} \left[\ln \frac{\mathrm{p}\left(X_e^{(i)}, X_t^{(i)}, X_v^{(i)}, agency^{(i)}\right)}{\mathrm{q}\left(X_t^{(i)}\right) \mathrm{q}\left(agency^{(i)}\right)} + \ln \frac{\mathrm{p}\left(\Theta\right)}{\mathrm{q}\left(\Theta\right)}\right]. \quad (8)$$

This bound can be computed using the joint probability from (6) and the conjugate priors from (7). Given that solution, calculating the update equations for the variational posterior $\mathrm{q}(agency)$ and the posterior hyperparameters $(\alpha, \beta, \kappa, \lambda)$ gives all necessary equations for the variational EM algorithm.

Computation of Agency Posterior

From a psychological viewpoint, it is interesting to investigate the posterior distribution of *agency*, specifically $p(agency{=}self|X_v)$, as it describes the probability that the subject interprets the visual stimulus as having been caused by the subject's own actions. Applying the sum and product rules of probability, we find

$$\mathrm{p}\left(agency^{(i)}{=}self|X_v^{(i)}, \mu_t^{(i)}, \Theta\right) = \frac{P_{self}\sigma_0 \exp\left(-\frac{1}{2(\sigma_t^2+\sigma_v^2)}\left(X_v^{(i)} - \mu_t^{(i)}\right)^2\right)}{\left(1 - P_{self}\right)\sqrt{\sigma_t^2 + \sigma_v^2}\exp\left(-\frac{1}{2\sigma_0^2}\left(X_v^{(i)} - \mu_0\right)^2\right) + P_{self}\sigma_0 \exp\left(-\frac{1}{2(\sigma_t^2+\sigma_v^2)}\left(X_v^{(i)} - \mu_t^{(i)}\right)^2\right)} \quad (9)$$

Results (1)

10 healthy subjects participated in the experiment. The parameters Θ were estimated using maximum likelihood optimization. To compare the assumption of a common set of parameters for subjects and the assumption of individual parameters for each subject, parameter estimation was performed both on the joint dataset over all subjects, as well as on individual subject's data.

	All Subjects	Subject 1	Subject 2	Subject 3	Subject 4
Model 1	**6347.32**	**566.21**	**645.14**	**608.69**	**530.76**
Model 2	6362.88	571.43	646.27	611.72	532.49

Table 1) Negative of maximized log-likelihoods.

Table 1) shows the negative of the maximized log-likelihood for all participants and single representative subjects. Both models accurately predict the data (Table 1), showing that small deviations between predicted and actual visual information were attributed to one's own action. This was not the case for large deviations, where subjects relied more on internal information.

Presented below are the results of representative subjects. The upper row holds the predictions of model 1 for the subjectively experienced sensory consequences of the subjects' motor actions. In the bottom row are the posteriors of agency attribution of the sensory feedback to oneself as predicted by model 1.

A clear distinction between the subjects' predicted agency posteriors can be seen in Figures 4) - 7).

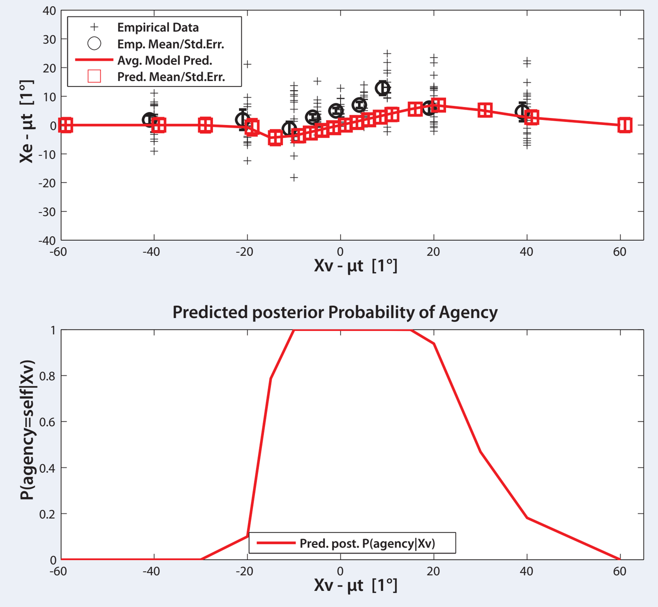

Fig. 4) Subject 1.

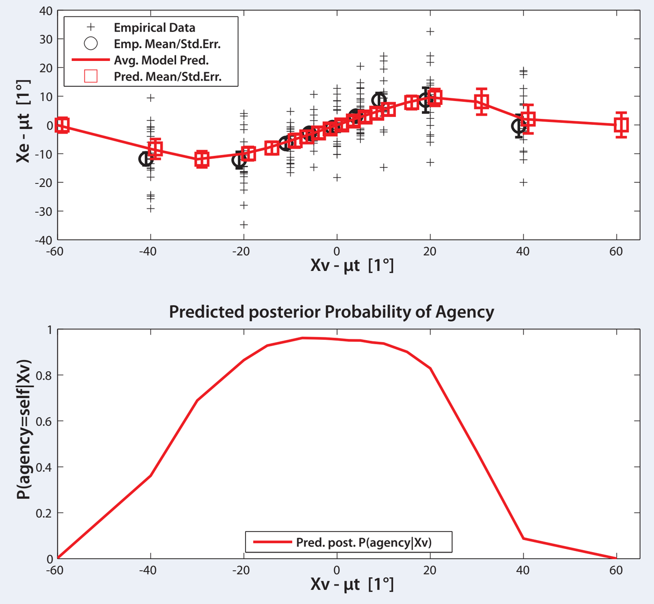

Fig. 5) Subject 2.

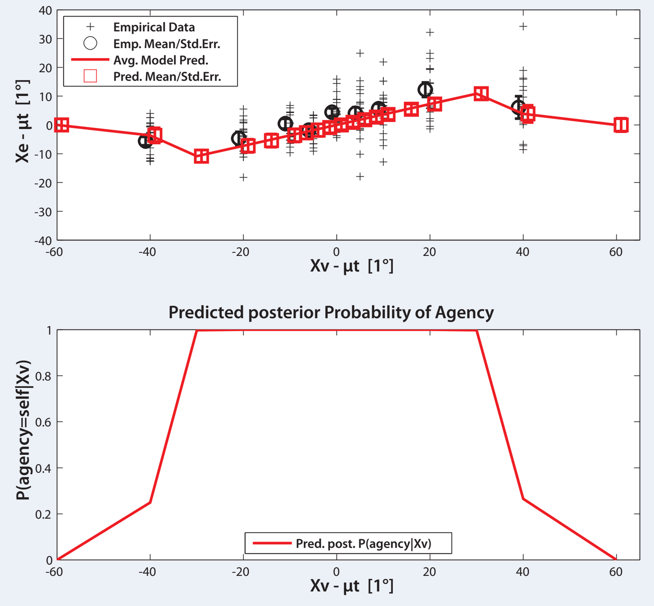

Fig. 6) Subject 3.

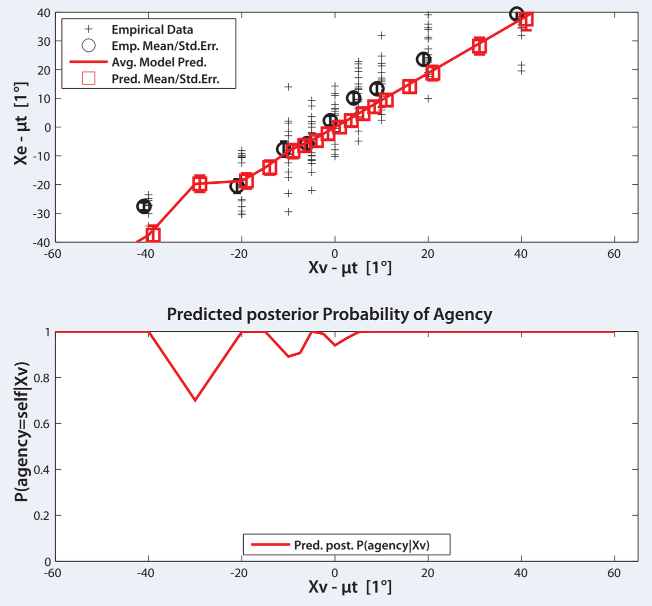

Fig. 7) Subject 4.

Results (2)

Fig. 4) shows a narrow agency posterior, signifying a sensitivity to deviations between internally predicted and actual sensory consequences. This is also evident in the difference between reported subjectively experienced sensory consequence and true motor action being a flat curve with a very small linear portion, i.e. the subjectively experienced sensory consequence were influenced more strongly by the internal motor state estimate than by the visual feedback.

Fig. 6) shows a wide agency posterior, on the other hand, implying a less stringent attribution of agency of sensory consequences to one's own actions. A stronger influence of the visual feedback on the subjectively experienced sensory consequence is visible in a more pronounced linear part, compared to Subject 1. Subject 2's agency posterior lies between those of Subject 1 and 3 (Fig. 5)).

Fig. 7) shows a case where the subject attributed agency of the visual feedback as sensory consequences to own motor actions, even for large deviations between visual feedback and true motor actions. This signifies a very low influence of the internal estimates, i.e. the proprioceptive measurements and the predictions by the internal forward models, on the subjectively experienced sensory consequences of one's actions compared to the visual stimuli.

Discussion and Conclusion

Subjectively Experienced Consequence. The tendency to attribute observed visual consequences to one's own action varied from subject to subject. For small deviations between real and predicted visual consequences of the own action, internal state estimate and visual feedback were integrated in a way that was similar to a linear cue fusion model (e.g. [7]). For large deviations, participants' subjective direction estimates were largely independent from the visual feedback.

Attribution of Agency. Consistent with this observation, we found a systematic variation of the posterior distributions of the agency variable with the size of the deviation: small deviations resulting in high probability, and large deviations resulting in low probability of attribution of agency for sensory consequences of motor actions to one's own actions. Subjects showed individually-specific tendencies (priors) for the attibution of agency and different regimes of deviations between predicted and actual visual feedback that resulted in linear fusion and the attribution of the visual stimulus to consequences of the own action.

Schizophrenia Patients. In some cases we found even for all conditions an attribution of agency to oneself as the cause of perceived sensory input. A similar effect was reported in adaptation experiments with schizophrenia patients [8], who, compared to healthy controls, showed a tendency to attribute perceived visual stimuli more to consequences of their own actions than to external influences.

Acknowledgements

This work was supported in part by the CIN Tübingen, the BCCN Tübingen (FKZ: 01GQ1002), the EU (FP7-ICT-215866 project SEARISE), the DFG and the Hermann and Lilly Schilling Foundation.

References

[1] D.M. Wolpert, Z. Ghahramani, M. Jordan. An internal model for sensorimotor integration. In *Science (1995), 269, 1880-1882.*

[2] K.P. Körding, U. Beierholm, W.J. Ma,S. Quartz, J.B. Tenenbaum, L. Shams. Causal Inference in Multisensory Perception. In *PLOSOne (2007), 2(9): e943.*

[3] L. Shams, U. Beierholm. Causal Inference in perception. In *Trends in Cognitive Sciences (2010), 14: 425-432.*

[4] D. Alais, D. Burr. The ventriloquist effect results from near-optimal cross-modal integration. In *Current Biology (2004), 14, 257-62.*

[5] J. Burge, M.O. Ernst, M.S. Banks. The statistical determinants of adaptation rate in human reaching. In *Journal of Vision (2008), 8(4:20):1-19.*

[6] C.M. Bishop. Pattern Recognition and Machine Learning. In *Springer Verlag, Berlin, (2007).*

[7] M.O. Ernst, M.S. Banks. Humans Integrate Visual and Haptic Information in a Statistically Optimal Fashion. In *Nature (2002), 415: 429-433.*

[8] M. Synofzik, P. Thier, D.T. Leube, P. Schlotterbeck, A. Lindner. Misattributions of agency in schizophrenia are based on imprecise predictions about the sensory consequences of one's actions. In *Brain (2010), 133: 262-271.*