Did I do it? 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

Introduction

The perception of own actions is affected by both visual information and predictions derived from internal forward models [1]. The integration of these 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]. Attribution of percepts to consequences of own actions should thus depend on the consistency between internally predicted and actual visual signals.
The 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.

Motivation

- Example: Archery.
- Aim: hit bullseye.
- 2 possible outcomes: Hit or Miss.

Fig. 1) Arrow shot.

- 'Hit': no change in action to achieve same successful result with next shot.
- 'Miss': correction to adapt one's action with next shot (for better result).
- Standard Bayesian approach both Hit and Miss → Fuse internal estimate and actual visual input.
- **BUT**: Error could be caused by (unpredictable or random) external influences, e.g. a sudden gust of wind.
- 'Miss by External Influence': attribute error to external influence, NOT wrong action execution → No cue fusion, no adaptation of action with next shot.

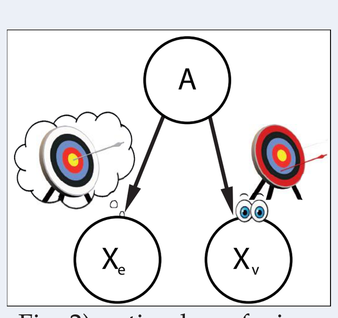


Fig. 2) optimal cue fusion.

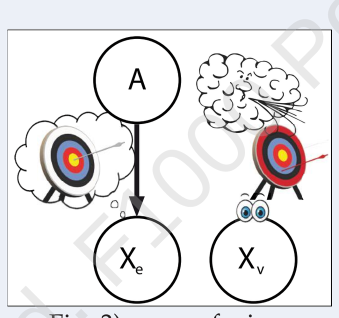


Fig. 3) no cue fusion.

- Models given for two cases of attributed agency of visual stimulus to self-action (Fig. 2) or to external influence (Fig. 3).
- Formulate causal inference model [2, 3], Fig. 4.

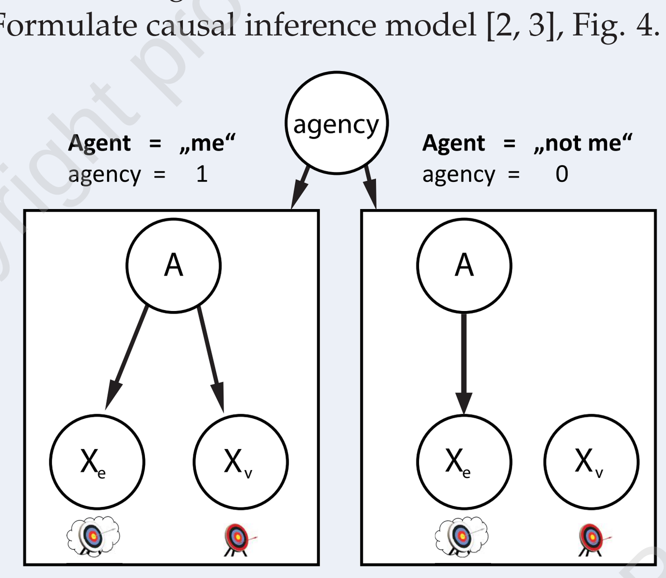


Fig. 4) Causal inference model [2, 3].

- Reformulate and adapt Fig. 4 to the experiment to receive the graphical model in Fig. 8.

Experiment

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.

- Tip of index finger of participant's right hand positioned on a haptic marker (trial starting position) on horizontal board.
- Hand invisible from participant's view.
- Straight, fast (quasi-ballistic) pointing movements to fixed target amplitude.
- No explicit visual targets.
- Target points chosen uniformly by subject within the upper right quadrant of the circle.

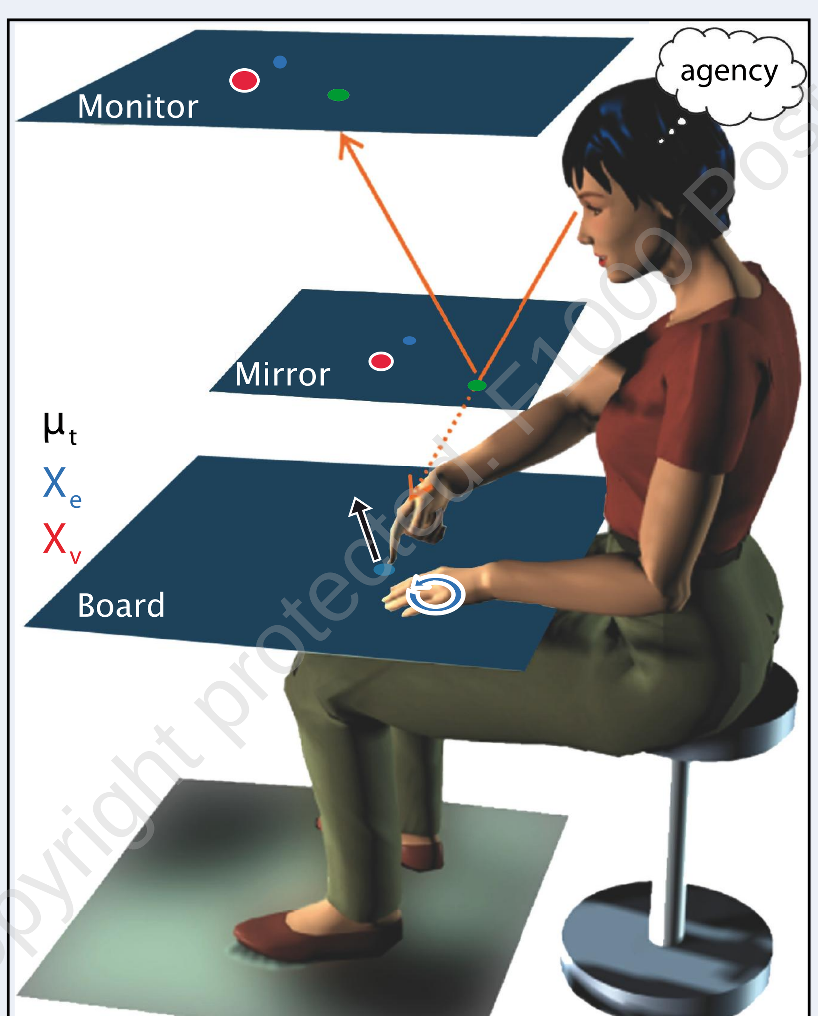


Fig. 5) Experimental setup.

- Terminal visual feedback $\mathbf{X_v}$ of motion at target amplitude.
- Visual feedback $\mathbf{X_v}$ calculated from tracked hand motion μ_t.
- Visual feedback either true or rotated by offset angle (i.e. 0°,±7°, ±14°, ±28°, ±56°).
- Offset angles in random order to minimize effects of trial-by-trial adaptation.

Experiment contd.

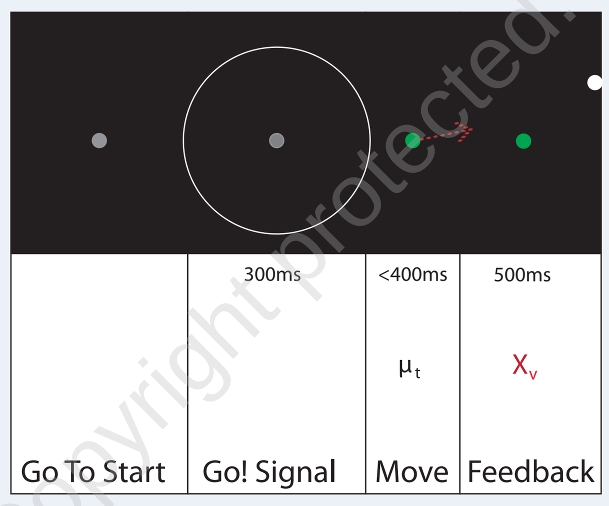


Fig. 6) Experimental paradigm, action and feedback.

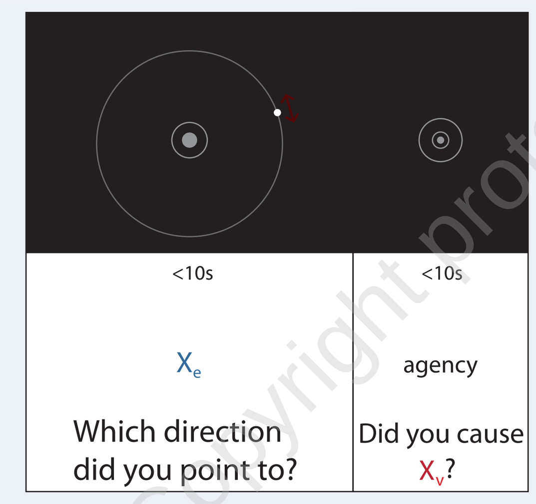
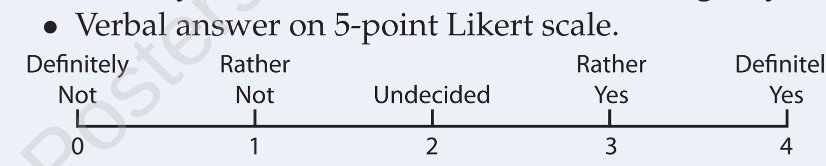


Fig. 7) Experimental paradigm, questions.

- Participants had to answer 2 questions:
- 1. **Which direction did you point to?** → $\mathbf{X_e}$
- Rotate controller with left hand to point cursor in desired direction.
- 2. **Did you cause the direction of $\mathbf{X_v}$?** → **agency**
- Verbal answer on 5-point Likert scale.



Graphical Model

- Binary gating variable (*agency*) modeling sense of 'agency'.
- **Fig. 8** both visual feedback $\mathbf{X_v}$ and internal motor state estimate $\mathbf{X_e}$ directly caused by the (unobserved) intended motor state $\mathbf{X_t}$.
- μ_t true motor action measured by the tracking system.

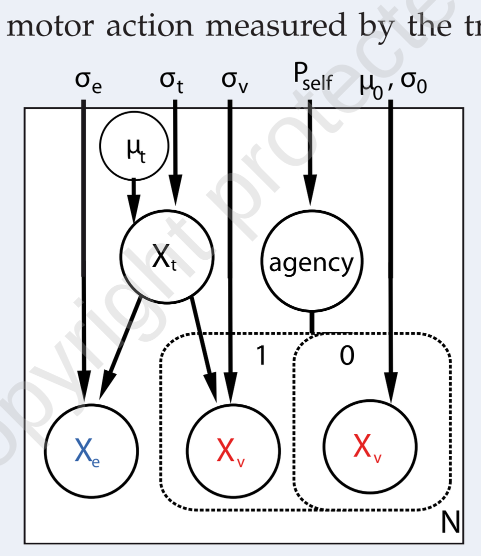


Fig. 8) Graphical model.

- From subject's perspective only observed variables: $\mathbf{X_v}$ and $\mathbf{X_e}$.
- Parameters to be estimated are the prior statistics for each variable:

$$
\begin{aligned}
X_t \sim \mathcal{N}\left(X_t|\mu_t,\sigma_t^2\right);\ X_e &\sim \mathcal{N}\left(X_e|X_t,\sigma_e^2\right)\\
X_v|\mathit{agency}{=}\mathit{self} &\sim \mathcal{N}\left(X_v|X_t,\sigma_v^2\right)\\
X_v|\mathit{agency}{=}\mathit{other} &\sim \mathcal{N}\left(X_v|\mu_0,\sigma_0^2\right)\\
\mathit{agency} &\sim \mathrm{Ber}\left(\mathit{agency}|P_{\mathit{self}}\right)
\end{aligned}
\tag{1}
$$

- Parameters learned over dataset $\{\mu_t^{(i)}, X_v^{(i)}, X_e^{(i)}\}$.
- Given these parameters, model predicts *agency* posterior.
- Prediction evaluated by comparison with measured agency ratings.

Maximum Likelihood Parameter Estimation

- Learn $\Theta = \{P_{\mathit{self}}, \sigma_e^2, \sigma_v^2, \sigma_t^2, \sigma_0^2, \mu_0\}$ by maximizing log-likelihood: $\sum_i^N \log \mathrm{p}\left(X_e^{(i)}|X_v^{(i)},\mu_t^{(i)},\Theta\right)$.
- Exploit factorization described by the graphical model (Fig. 8), giving Eq. (2):

$$\mathrm{p}\left(X_t^{(i)}, X_e^{(i)}, X_v^{(i)}, \mathit{agency}^{(i)}|\mu_t^{(i)},\Theta\right) = \mathrm{p}\left(X_t^{(i)}\right)\mathrm{p}\left(X_e^{(i)}|X_t^{(i)}\right)\mathrm{p}\left(X_v^{(i)}|X_t^{(i)},\mathit{agency}^{(i)}\right)\mathrm{p}\left(\mathit{agency}^{(i)}\right). \tag{2}$$

Computation of Agency Posterior

- Psychological viewpoint: interesting to investigate posterior distribution of *agency*.
- Eq. (3) describes probability that subject interprets visual stimulus as caused by own actions.

$$\mathrm{p}\left(\mathit{agency}^{(i)}{=}\mathit{self}|X_v^{(i)},\mu_t^{(i)},\Theta\right) = \frac{1}{\frac{(1-P_{\mathit{self}})\sqrt{\sigma_t^2+\sigma_v^2}}{P_{\mathit{self}}\sigma_0}\exp\left(\frac{1}{2(\sigma_t^2+\sigma_v^2)}\left(X_v^{(i)}-\mu_t^{(i)}\right)^2 - \frac{1}{2\sigma_0^2}\left(X_v^{(i)}-\mu_0\right)^2\right)+1}. \tag{3}$$

Results

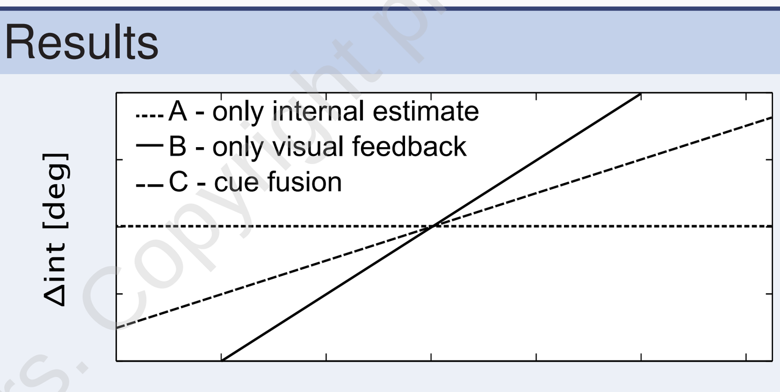


Fig. 9) Expected strategies along which data should cluster.

- $\Delta\mathrm{vis} = \mathbf{X_v} - \boldsymbol{\mu_t}$; $\Delta\mathrm{int} = \mathbf{X_e} - \boldsymbol{\mu_t}$

- **Fig. 9** shows components along which we expect data to fall for different strategies.
- If subjects trust only their internal estimate, we expect data to accumulate around $\Delta\mathrm{int} = 0$, (A).
- If, on the other hand, subjects trust only the visual feedback, the data is expected to accumulate around $\Delta\mathrm{int} = \Delta\mathrm{vis}$, (B).

Region coloring scheme.

6 healthy subjects participated in the experiment, representative subjects are shown below.

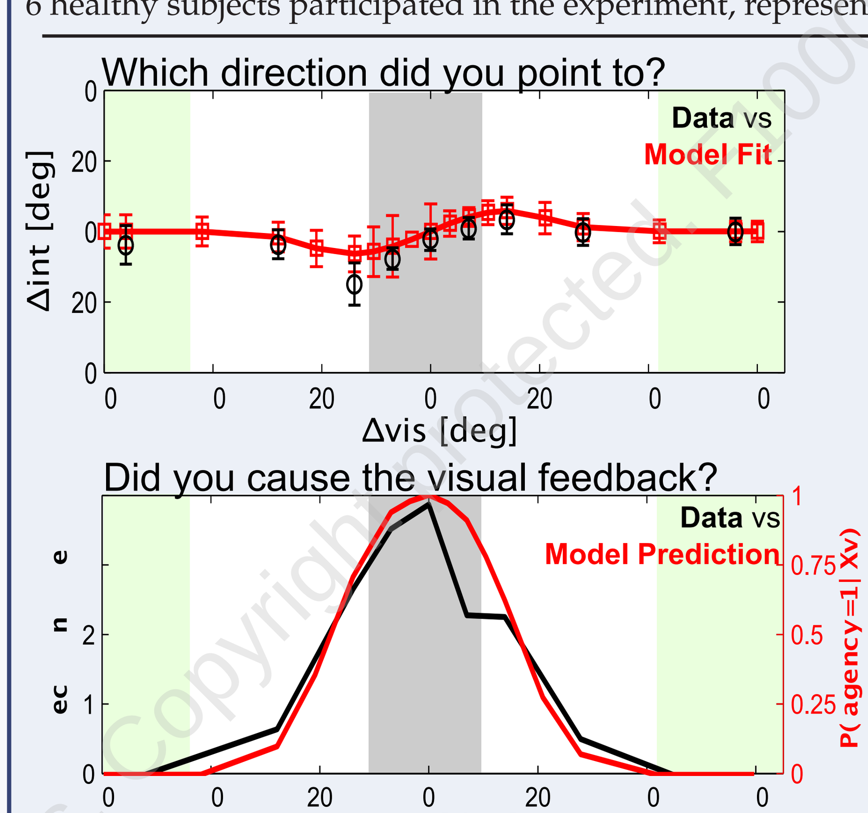


Fig. 10) Subject 1.

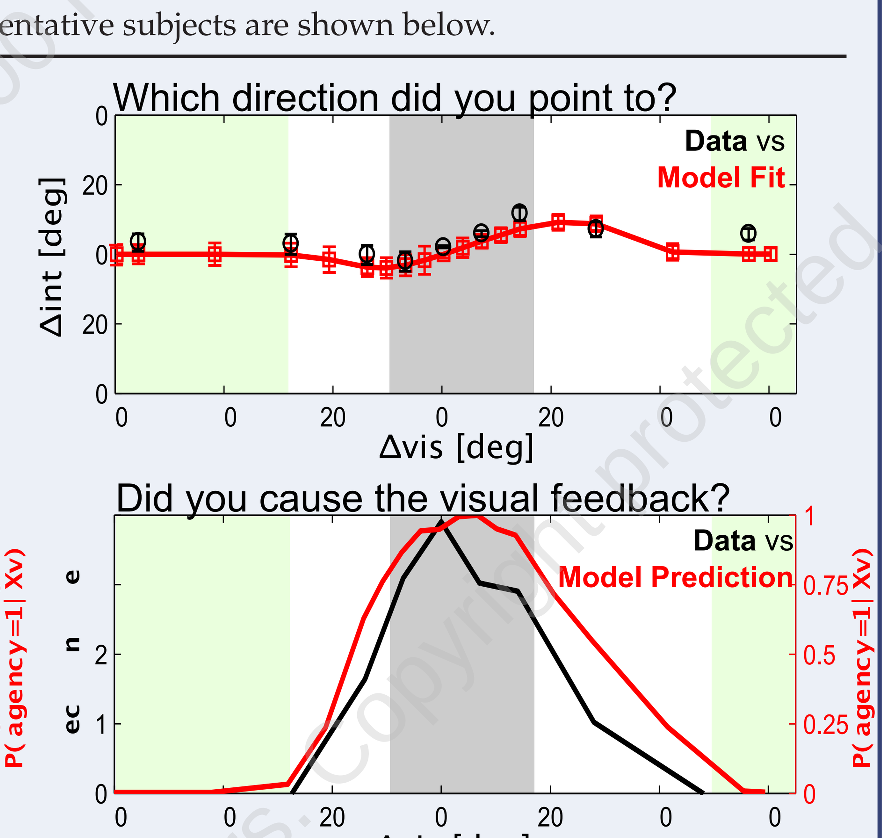


Fig. 11) Subject 2.

- Different regions with respect to Δvis can be identified in the data and model fit:
- Outer regions: no influence of $\mathbf{X_v}$ → no cue fusion.
- Inner region: strong influence of $\mathbf{X_v}$ → cue fusion.
- Subject specific bias visible to over- or underestimate direction $\mathbf{X_e}$.
- Identified regions match the subjective attribution of agency of $\mathbf{X_v}$ as consequence of self-action.
- Model predicts subject's posterior agency attribution correctly.

Subject	Q^2	Q^1	Width ratio (Pred/Data)
1	97%	83%	99%
2	96%	80%	123%
3	96%	82%	127%
4	79%	72%	166%
5	78%	48%	181%
6	73%	41%	187%

Table 1) Goodness of Fit of agency posterior.

$d(\Delta)$ = normalized emp. data of agency posterior
$p(\Delta)$ = normalized prediction of agency post.

$$Q^2 = 1 - \frac{\int (d(\Delta)-p(\Delta))^2\,\mathrm{d}\Delta}{\int d(\Delta)^2\,\mathrm{d}\Delta} \tag{4}$$

$$Q^1 = 1 - \frac{\int (d(\Delta)-p(\Delta))\,\mathrm{d}\Delta}{\int d(\Delta)\,\mathrm{d}\Delta}. \tag{5}$$

- Quantitative analysis of the prediction quality confirms the first visual impression.
- Correct prediction of agency posterior.
- Width ratio shows a trend of the model to overestimate the cue fusion region.
- Model's tendency to over-attribute agency also reflected in lower Q^2 and Q^1 scores.

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.*

Previous Exp. Results

10 healthy subjects participated in the previous experiment, representative subjects are shown below.

- Different Manipulation angles (i.e. 0°,±5°, ±10°, ±20°, ±40°).
- Only: **Which direction did you point to?** → $\mathbf{X_e}$

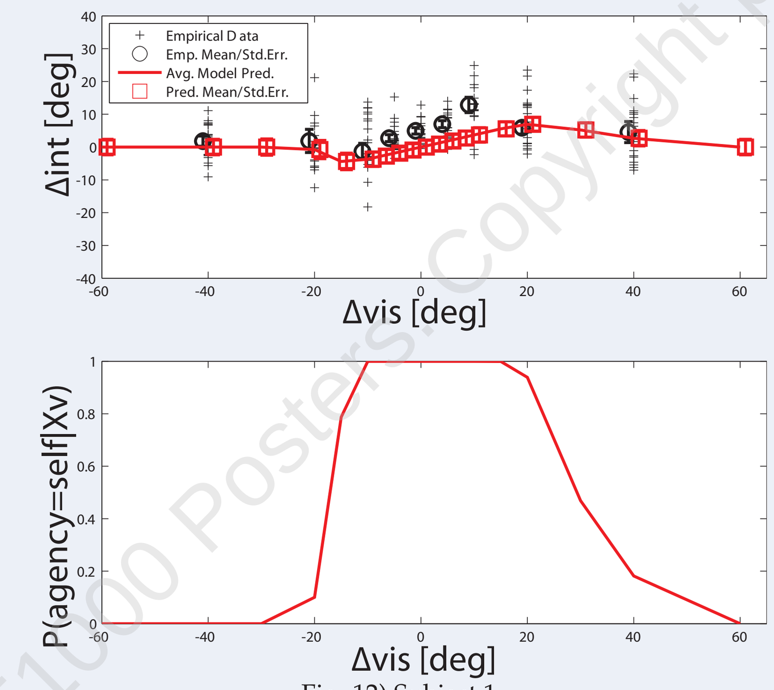


Fig. 12) Subject 1.

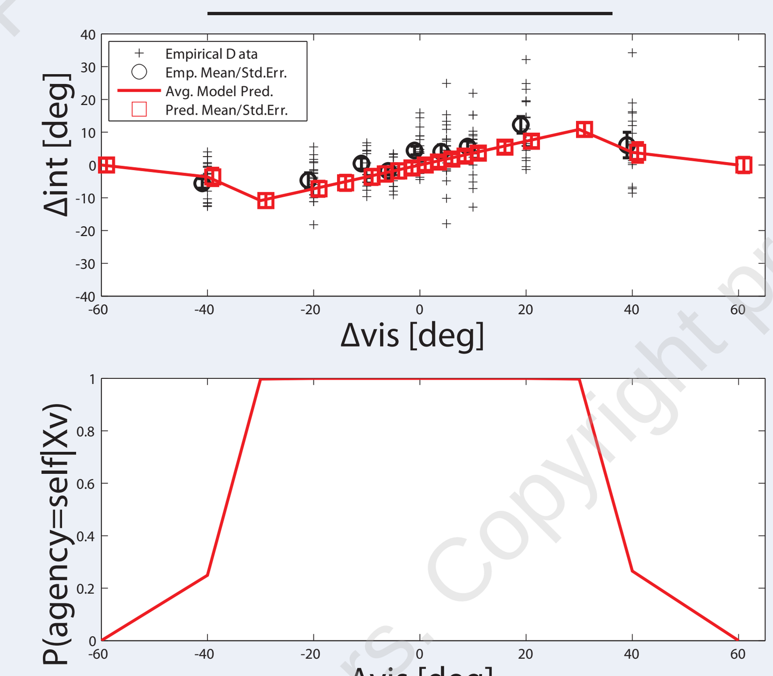


Fig. 14) Subject 2.

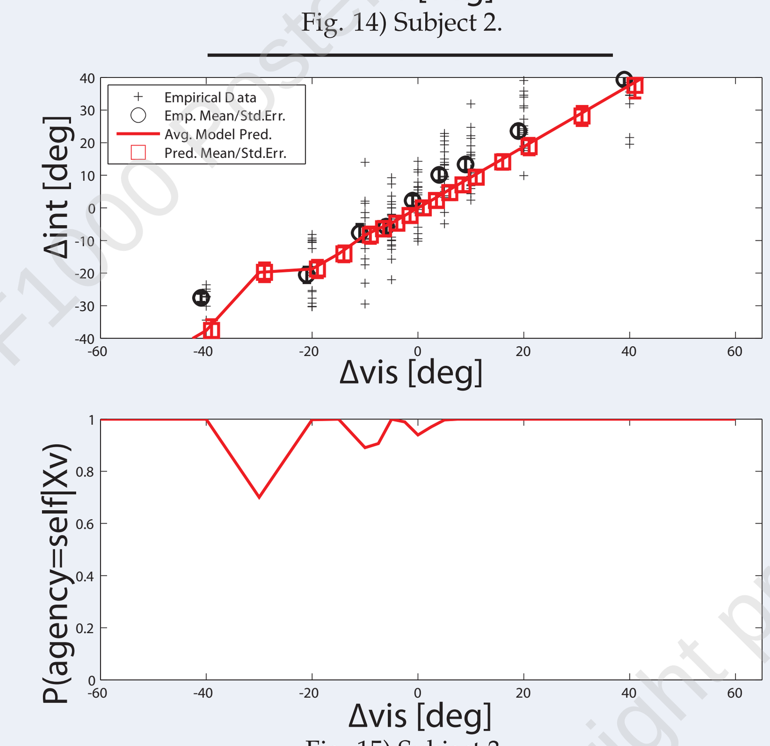


Fig. 15) Subject 3.

- Subject to subject variations in agency attribution tendency.
- **Fig. 12** narrow agency posterior, signifying a high sensitivity to deviations.
- **Fig. 14** wide agency posterior, attribution of agency of sensory consequences to one's own actions. Stronger influence of visual feedback on the subjectively experienced sensory consequence.
- **Fig. 15** attribution of agency of visual feedback as consequence to own motor action, even for large deviations.

Discussion and Conclusion

Subjectively experienced consequence.

- Small Δvis (deviation btw. real and predicted visual consequences) → optimal fusion of internal estimate and visual feedback.
- Large Δvis → direction estimates being largely independent from visual feedback.
- Good fit of model expectation to data, $\mathbf{X_e}$, though no bias.

Attribution of agency.

- Systematic variation of *agency* posterior with the deviation size.
- Small Δvis → high probability.
- Large Δvis → low probability.
- Correct prediction of 'agency' (belief that observer caused the visual feedback) by this model (Fig. 8).

In conclusion,

- Subjects attribute agency of sensory consequences to their own motor actions depending on individual parameters.
- Δvis affects agency attribution.
- Optimal cue fusion is performed within region of self-attribution and not outside.
- Bayesian Graphical Model suitable to capture underlying model selection.
- The presented model correctly predicted the agency posterior.

Acknowledgements

This work was supported in part by the Bernstein Center for Computational Neuroscience, Tübingen, funded by the German Federal Ministry of Education and Research (BMBF; FKZ: 01GQ1002), the CIN Tübingen, the EU projects FP7-ICT-215866 SEARISE, FP7-249858-TP3 TANGO, FP7-ICT-248311 AMARSi, the DFG and the Hermann and Lilly Schilling Foundation.