The Robo-havioral Methodology: Developing Neuroscience Theories with FOES

Maria Otworowska[*ab] Jordi Riemens[*a] Chris Kamphuis[a]

Pieter Wolfert[a] Louis Vuurpijl[a] Johan Kwisthout[†ab]

[a] Department of Artificial Intelligence
[b] Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen, PO Box 9104, 6500HE Nijmegen, The Netherlands

Abstract

In systems neuroscience there is a big gap between what theorists postulate (i.e., grand unifying theories about the general principles underlying cortical processes such as the predictive processing account) and what empiricists measure (i.e., reaction times, pupil dilations, blood-oxygenated level dependent signal in brain areas, magnetic pulses). It is becoming increasingly difficult for the theorists to come up with empirically testable hypotheses and for the empiricists to use their findings to confirm or refute a theory. We propose a research methodology based on robot simulations that may help bridge that gap. The methodology is summarized by four keywords: *Formalize* verbal theories into computational models; *Operationalize* this computational model into a working robot implementation; *Explore* the consequences of various design choices and parameter settings to generate empirically testable hypotheses; and finally *Study* these hypotheses in behavioral or imaging experiments. We lay out a research program that aims at investigating various open issues in predictive processing and exemplify our approach in a simple case study.

1 Introduction

The *predictive processing* account [1, 5] is currently one of the most influential unifying neuroscientific accounts of what drives our cortical processes. According to this account, the brain continuously makes predictions about what inputs it will get, based on hierarchical (stochastic) generative models that maintain the current best hypotheses of the causes of these predictions; updating the models based on precision-weighted prediction errors. It is claimed that this prediction principle applies to the entire cortex and that the same broad apparatus and mechanism is used for both lower and higher cognition, e.g., both low-level vision and intention attribution [1]. While the predictive processing account is fleshed out computationally for low-level vision and motor control [3], to account also for "*higher cognitive phenomena such as thought, imagery, language, social cognition, and decision-making*" there is still "[...] *plenty of work to do.*" [5, p.5]. Our research groups have been contributing to this research program both by empirical research [6,17] and by theoretical and computational contributions [4,8–11,14,20,21]. One of our key theoretical contributions is in the formalization of predictive processing in terms of causal Bayesian networks [15], thus allowing for the representation of structurally rich knowledge domains and complex interactions [8, 9].

Ideally, there is an interplay between theoretical advances and empirical research, where the theoretical work suggests hypotheses to test and the empirical work updates the theory. In current research on predictive processing, however, we experience a huge gap between theoretical and empirical work,

∗ Both authors share first authorship.
† Corresponding author: j.kwisthout@donders.ru.nl.

giving rise to many methodological and foundational questions. In particular since the predictive processing account offers a *unifying* view on cortical processes, often painting with rather broad strokes, it is difficult to determine whether a concrete empirical result really is evidence in support of predictive processing; as the theory is so general, almost every result can be explained to match the general principles of the predictive processing account [16]. It is often questionable whether results and explanations are 'real' or whether they are mere idiosyncrasies of an incorrect, incomplete, ambiguous, or ill-defined interpretation of the (verbal) theory or the experimental manipulation thereof; this problem comes on top of the surprisingly bad reproducibility [2] of experiments. The following example may illustrate our concerns.

Model revision or model selection?

Sometimes prediction errors indicate that our current generative models need to be revised. For example, if we observe our vegetarian friend ordering a ham sandwich, chances are that our friend decided to quit being a vegetarian. That event impacts our future expectations. However, if she was not a vegetarian, but we knew that she just ordered cheese sandwiches more often than ham sandwiches, the observation perfectly fits with our model and there is no reason for revision. To put this in terms of predictive processing: in the first example the precision of the prediction error is higher, leading to an increased tendency for model revision, whereas in the second example the prediction error is damped due to the expected uncertainty in the prediction. O'Reilly et al. tried to experimentally dissociate different effects of prediction errors, based on the amount of expected uncertainty in the predictions [13]. In a saccadic planning task, participants focused on dots appearing in a particular quadrant of a unit circle, where sometimes dots appeared in other quadrants. Whether this indicated a 'block change' (i.e., the quadrant of interest was changed) or just a one-off trial (the quadrant of interest remained the same) was indicated by the color of the outlier dot. The investigators found different brain patterns and reaction times for the different conditions and suggested distinct brain networks associated with 'surprise' and 'updating' processes. It is debatable, however, how to fit these results in the predictive processing theory. One explanation may be that the results are evidence that the brain revises its generative models based on unexpected prediction errors, that is, that it "*need[s] to update the internal model to predict future observations accurately in a changeable environment*" [13, p.E3660]. An alternative and perfectly fitting explanation is that the different color acts as a contextual cue that leads to the selection of a different sub-model that now becomes active in generating hypotheses. Whereas it is plausible that models are generated and revised during the learning stages of the experiment, it is quite defendable that what the researchers measured is model *selection* based on context, not model *revision*.

A crucial assumption in the predictive processing account is that the brain tries to predict its inputs based on previously developed generative models; it processes only that part of the input that is yet unexplained (i.e., that was not predicted). To stick with the previous example, our brain may have built up a generative model of what we expect our friend to order (cheese sandwiches) at lunch time; only if there is a prediction error—a ham sandwich, rather than a cheese sandwich—the conflicting part of the input (ham instead of cheese) is processed by the brain. There are multiple ways how such prediction errors can be dealt with [8]. Depending on the context of the situation and on the confidence we have in the model, we can potentially lower prediction error by either revising our generative models, updating our current beliefs without changing the model, obtaining additional observations, or by actively manipulating the inputs. Respective examples of these mechanisms are to update the model of our friend's dietary preferences when we see her ordering a ham sandwich, updating the belief *either heads or tails with uniform probability* to *definitely heads* on observing the outcome of a coin toss, determining whether either our own train or the train on the parallel track starts to move by looking at the (stationary) railway station building, or scratching one's head to get rid of an unwanted itch.

Currently, the predictive processing account is largely silent as to which mechanism is applied under which circumstances [9]. As the running example illustrates it is far from straightforward to empirically disentangle these candidate mechanisms. The reasons therefore are two-fold: On the one side, experimentally controlled manipulations will be based on ambiguous and incomplete verbal theories. It is next to impossible to decide whether a particular empirical finding captures a 'real effect' or whether it is due to implicit 'design choices' in the theory that emerge from the particular experimental setup. On the other side, even if we are confident that a particular finding captures a 'real effect', the gap between verbal theory and experimental condition hinders the use of that finding as confirmation or refutation of a particular aspect of the theory. To help overcome both issues we propose the *Robo-havioral methodology* as an intermediate step in the scientific process. The goal of this paper is two-fold. Firstly (Section 2), to flesh out this methodology in more detail and to propose a research program built on it; secondly (Section 3), to present a small case study that nicely illustrates how even a relatively simple implementation can already reveal structural gaps in the theory and how we propose to deal with them. We conclude in Section 4.

2 The Robo-havioral Methodology

Inspired by a longstanding tradition in Artificial Intelligence (already advocated in [12], but see also approaches such as [18]), we propose a new research methodology, somewhat tongue-in-cheek called the *Robo-havioral methodology*, as an intermediate step between theoretical and empirical investigations using so-called *FOES*. Here, FOES is the abbreviation of *formalize, operationalize, explore, and study*: transform the verbal theory into a formal computational model, operationalize the model into a working robot, explore the 'parameter space' and 'design considerations' of the theory in robot simulations, and study thus obtained hypotheses in behavioral or imaging experiments; all in a continuous cycle. We need to emphasize that the goal of this methodology is *not to build smart robots* or to extend the state of the art in robotics. In contrast, the goal is to find gaps and ambiguities in neuroscientific theories, identify 'design choices' in such theories, and to explore the consequences of such commitments. The operationalization into working robots forces us to be complete and consistent with respect to our theoretical commitments, thereby improving the theory (understanding-by-synthesis). The exploration allows us to go beyond thought experiments and purely computational modeling; it generates empirically testable (i.e., with human participants) hypotheses that naturally follow from such design considerations and parameter settings. They are grounded on a full, complete, and consistent operational model, rather than on the scientist's speculations. This helps to ensure that whatever experimental results we obtain can be related to the theory. To summarize, we propose that this methodology forms a *trait d'union* between theory forming and experimenting, leading to the following picture.

1. **Conceptual analysis and theory forming**
   To propose verbal theories that potentially explain phenomena of interest.
2. **Robo-havioral exploratory studies**
   To explore, identify, and fill the gaps in the theories and generate testable hypotheses.
3. **Behavioral and imaging experiments**
   To empirically test the hypotheses and refine the theories and models.

Computer simulations can be very worthwhile tools for exploration and hypothesis generation. We do make extensive use of our *predictive processing toolbox* to compare different scenario's etcetera. However, we propose to operationalize theories into (embodied, embedded, task-oriented) robots for the simple reason that this forces us to take reality into account. While by formalizing a verbal theory into computational models already quite some ambiguities typically can be resolved, the proof of the pudding really is in the eating—not in writing up the recipe. To cite Barbara Webb [19, p. 1084]: "*Although some simulations may include all the environmental details of the real world, the simple fact is that the majority of simulations do not. Rather, they include what the modeler thinks to be important, that is, they tend to be biased towards the hypotheses to be tested.*" In particular since one of the goals of this methodology is to find ambiguities and missing details in theories, we believe it is vital to have the operationalizations act in the real world.

3 Case study: Recognize Sorting Intentions

Many aspects of the predictive processing account are still under-defined, in particular when we move to higher cognitive theories such as action understanding [7]. For example, we do not yet know how the formal theory can account for the integration of long-term intentions with motor acts on a much shorter time-scale, how generative models are developed, what the effect of different mechanisms of prediction error minimization is, etcetera [9]. In addition, many such aspects are probably still overlooked. To illustrate our methodology, we constructed a LEGO Mindstorms RECOGNIZESORTINGINTENTIONS robot whose goal it is to recognize whether someone is sorting Duplo blocks either by color or by size (Figure 1). There are nine blocks, three colors (red, green, yellow) and three sizes ($2 \times 2 \times 4$, $2 \times 2 \times 2$, $1 \times 2 \times 4$). The blocks thus can be sorted by the three colors or by the three sizes. When picking up the blocks from a stack, and putting them into three separate bins, initially it is not clear what the sorting intention is, but at most after the fourth block is put in one of the three bins, this can be inferred. The robot is equipped with various sensors and it can move around to change its sensory inputs. Its "action understanding" is based on a hierarchical generative model, i.e., a hierarchy of Bayesian networks that generate predictions, at the lowest level (sensory) predictions for the inputs of its sensors, at higher levels (action) predictions for movements towards buckets and the grasping of blocks, predictions for where a particular block is put, which block is selected to pick up, and at the highest level (intention) predictions about which sorting strategy is used. The robot will encounter uncertainty (due to noisy sensors or ambiguous situations) and prediction errors; we implement and compare means of dealing with them (for example, moving closer or turning its head to increase precision of the sensors (Figure 2) or get additional observations, changing beliefs of the actual sorting strategy, etc.).

Figure 1: Sorting Duplo blocks either by color (top) or by form (bottom). We construct a robot, based on the predictive processing account, that aims to recognize which sorting intention is applied. Note that in principle it suffices to observe four blocks at most to be able to disambiguate the sorting intention if there are 3×3 blocks and 3 bins. For example, if the three bins contain a red 2×2, a green 2×4, and a yellow 1×4 block, respectively, any subsequent block disambiguates the sorting intention.

This simple setup is already quite rich, allowing for many studies. For example, it allows for studying how a robot can learn causal relationships, how different strategies of dealing with prediction errors effect the agent's efficiency or learning capabilities, etcetera. However simple, the setup is already quite challenging, in particular with respect to the 'lower-level capabilities' of getting to know which block is held at a time and where it is going to. However, the focus of our methodology is not on this level. We are not as much interested in aspects of computer vision and object recognition, but rather in higher-level ways of dealing with uncertainty. For the concrete case study that we report on in this paper, we have (from the robot's perspective) discrete time steps, discrete block movement, and fixed locations of the blocks. This very implementation of such a highly constrained case study already proved very valuable in explicating overlooked aspects of the predictive processing account (e.g., that reducing prediction error by acting (relocating the position of its sensors) needs counter-factual models of how actions potentially change perception) and hinted at to many interesting follow-up research questions. In the next sub-section we will describe the computational (Bayesian) model of this case study, highlighting where the process of formalization, operationalization and exploration exposed fundamental or practical issues. We emphasize again that our goal was not to build a smart robot for recognizing Duplo blocks sorting procedures, but to become aware of such issues, i.e., understanding-by-synthesis. The description of the computational model and the pilot will therefore be interleaved with "Theory Insight" comments.

Figure 2: An ambiguous percept that can be disambiguated by the robot by moving the position of its sensors. In the left panel, the robot cannot disambiguate between a 1×4 or a 2×4 Duplo block. When it moves its position (right panel), thereby changing the inputs of its sensors, it might reduce the prediction error stemming from this uncertainty.

3.1 Constructing the computational model

We focus our computational model on a time lapse of four blocks that are taken from a stack and placed in one out of three bins. The model is graphically depicted in Figure 3. At the highest level of the generative model (left panel) is a binary variable describing the sorting intention with values SORT-BY-COLOR and SORT-BY-FORM. This variable, together with contextual information (the blocks that are still on the stack and the contents of the bins), generates predictions for which block will be picked next from the stack, transported to one of the three bins, and placed in that bin. In the model this is called a COMPOUND-ACTION. Initially, the probability distribution of the first action will be uniform. This COMPOUND-ACTION can be further decomposed in more basic actions PICKUP-X, MOVE-X-TO-Y, and DROP-Y (right panel). Each of these actions drives predictions for what the robot will observe in any of the four discrete frames T_0 (when the block is picked up), T_1 (when the block hovers above the first bin), and T_2 and T_3 (when the block is above the second and third bin). For example, if we expect the red 2×2 block to be moved to bin 2, then we expect to see this block in T_0, in T_1, and in T_2, but we expect to see nothing in T_3. In addition there is a dynamic dependence: if we (think we) saw the red 2×2 block in T_0, then we don't expect a green 1×4 block in T_1. The predictions here finally drive low-level predictions for the color (C) and shape (S) expectations of the robot.

Theory Insight In the predictive processing account, information flows top-down (e.g., from hypothesized intentions to predicted actions) and bottom-up (from observed color and shape to action). However, it seems almost inevitable to have dependences in time as well (i.e., whatever block we saw one frame ago). The Bayesian inference computations must deal with such time-dependences, forbidding to (or at least be reluctant to) change belief states in the past. This aspect is overlooked in the predictive processing account.

Theory Insight Much general world knowledge (e.g., a block cannot be both on the stack and in one of the bins at the same time; blocks do not change color), needs to be encoded for predictions to take shape. We operationalized this in the case study as constraints encoded in the probability distributions. However, this becomes problematic if one wants to update probability distributions due to prediction errors: some beliefs (such as the laws of nature) should be less resilient than others.

So far, we described the top-down predictive stream in predictive processing. The hypothesized sorting intention (initially uniformly distributed) drives a prediction for the actions, and the hypothesized actions (again initially uniformly distributed) drives a prediction for the visual stimuli in each of the four frames. The PREDICTION process in the predictive processing account thus is simply the computation of a posterior distribution [8]. At the lowest level, the prediction is matched with the actual observation of color and shape by the robot. To obtain these observations, we equipped the robot with a web-cam and processed the raw inputs such as to obtain probability distributions over the three colors and the three shapes (and $\varnothing$ in both cases to indicate nothing was seen), as described in the supplementary material available at `www.socsci.ru.nl/johank/BNAIC2015/technicalreport.pdf`. We then computed the PREDICTION ERROR as the Kullback-Leibler divergence between the predicted distributions for color and shape and the observed distributions. The next step is to update the hypotheses as to reduce prediction errors. In this simple case study we assume that prediction errors are always reduced by updating current beliefs. This BELIEF REVISION process seeks to find a revised prior probability distribution over the set of hypothesis variables that minimizes the prediction error.

Theory Insight Color and shape can be seen as multi-modal predictions. We thus have, for a particular hypothesis, *two* predictions and *two* prediction errors. We may or may not *weight* the separate prediction errors according to how influential we want them to be in updating our beliefs. This may or may not be orthogonal to the confidence in (a particular part of) the generative model.

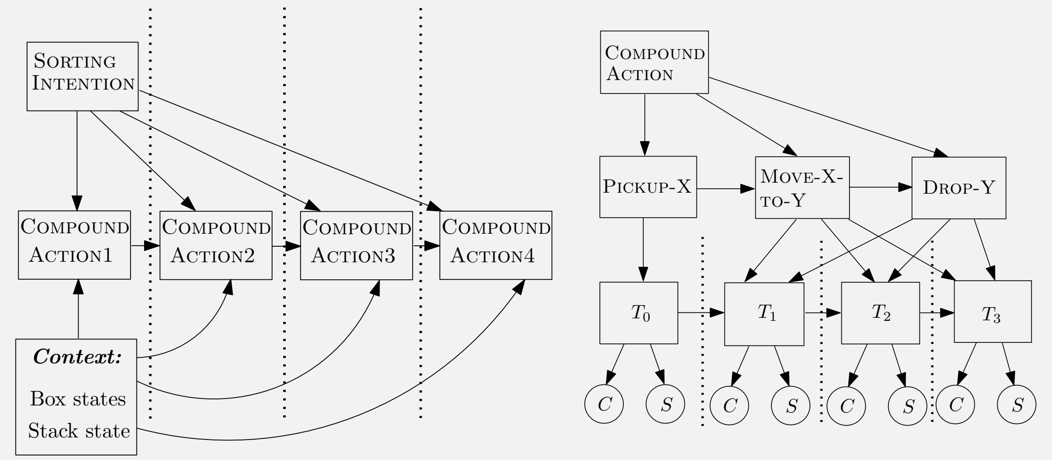

Figure 3: Description of the computational model of the RECOGNIZESORTINGINTENTIONS robot. The model consists of hierarchical layers, describing intentions, context-modulated actions (left panel), visual percepts at various points in time, and raw color and shape information (right panel). Every layer is described by a causal Bayesian network. Bottom-up inference is done by computing prediction errors between expected and actual inputs, and revising the probability distribution over the hypothesis nodes such that this prediction error is minimized. This again induces a prediction error at the level above which is then propagated upwards.

3.2 Experimental pilot

In the experimental setup four Duplo blocks (red 1×4, yellow 2×4, green 2×2, and red 2×4) were distributed among three bins in sixteen distinct time frames (four time frames for one block), in a following order: red 1x4 to bin 1, yellow 2x4 to bin 3, green 2x2 to bin 2 and red 2x4 to bin 3. This particular distribution of the blocks over the bins represents the sorting strategy SORT-BY-FORM. Initially, the robot has uniform predictions for all actions and observations, yielding the prediction $\{\Pr(red) = 0.5, \Pr(yellow) = 0.25, \Pr(green) = 0.25, \Pr(\varnothing) = 0\}$ for color, and $\{\Pr(2 \times 4) = 0.5, \Pr(2 \times 2) = 0.25, \Pr(2 \times 4) = 0.25, \Pr(\varnothing) = 0\}$ in T_0. According to the robot's sensors, in T_0, the robot actually perceived the following distribution for color: $\{\Pr(red) = 0.3719, \Pr(green) = 0.1344, \Pr(yellow) = 0.272, \Pr(\varnothing) = 0.2217\}$ and for shape: $\{\Pr(2 \times 4) = 0.3407, \Pr(2 \times 2) = 0.1418, \Pr(1 \times 4) = 0.2957, \Pr(\varnothing) = 0.2217\}$. This yielded a prediction error between predictions and observations of 0.1225 for color and 0.1266 for shape. The prediction errors were minimized by revising the probability distribution over the hypothesis variable T_0 to $\{\Pr(\text{red}\, 1 \times 4) = 0.3239, \Pr(\text{yellow}\, 2 \times 4) = 0.1631, \Pr(\text{green}\, 2 \times 2) = 0.1097, \Pr(\text{red}\, 2 \times 4) = 0.1631, \Pr(\varnothing) = 0.2399\}$, yielding a prediction error of 0.004 for color and 0.002 for shape. While there is considerable uncertainty, the robot still correctly inferred the red 1×4 block to be the most probable one in T_0. The now induced prediction error at the block level was propagated to the action level, giving an updated distribution for the COMPOUND-ACTION variable. However, our software had great difficulty to find an updated distribution for the highest level of the hierarchy, SORTING INTENTION. Upon closer inspection this was the case because all combinations of all contextual influences (the blocks still on the stack and in the bins) need to be taken into consideration in revising the distribution, giving too much degrees of freedom for tractable computations.

Theory Insight Whatever sorting intention is most likely depends on many bits of contextual information, such as the current state of the stack and the bins. When we represent such contextual information (and the implicit world knowledge) as a probability distribution, as we did in this case study, a straightforward belief revision algorithm will marginalize over all values of all these variables. A 'smarter' way of representing world knowledge and contextual information is needed for computations to be tractable.

Halting the inferential stream at the COMPOUND-ACTION variable and making again predictions for the observations in T_1 and T_2 and updating the probability distributions after prediction errors, the robot correctly predicted it will not see any block in T_3, i.e., the block was placed in bin 2. It had some difficulty in discriminating the $\mathrm{red}\,1 \times 4$ from the $\mathrm{red}\,2 \times 4$ block, however; as we could not further infer the updated probability distribution of the SORTING INTENTION due to the intractability issues.

3.3 Discussion

Even this simple case study is illuminating. Being forced to actually *implement* predictive processing enforced various design considerations upon us, such as how to encode world knowledge (e.g., a block does not suddenly change its shape and it cannot be in two bins at the same time) and how to deal with multi-modal prediction errors. The consequences are not trivial: for example, one of our assumptions (that world knowledge can be encoded in probability distributions, disallowing 'impossible' combinations of values by attributing a zero probability to them) gives rise to huge intractability issues when prediction errors need to be minimized. Our conclusion is that even in this simple pilot much information is gathered that enforces us to rethink parts of the verbal theory ("go back to the drawing board").

One could question whether it was really needed for this simple case study to construct a robot, i.e., whether a computational simulation would not have revealed similar insights. We agree that the full strength of this methodology will be more apparent in studies with a closed action-perception cycle, where the robots act upon the environment and influence their future perceptions; this early report does not reveal this full strength. However, even in this simple study there were theory insights (like the issue of multi-model prediction errors and how to weight them in updating models) that we believe we would have missed had we resorted to a computational simulation.

4 Conclusion

We proposed the Robo-havioral methodology for bridging the gap between theoretical and empirical research in neuroscience. The aim of this methodology is to a) identify weak spots in verbal theories and b) propose informed empirical hypotheses that are closely related to theoretical commitments. The methodology is based on the principle *understanding-by-synthesis* and can be paraphrased by four keywords: Formalize, Operationalize, Explore, and Study. We showed how even a simple case study can already expose various theoretical issues in the predictive processing account, simply by "trying to make it work in the real world", providing food for thought and subsequent experiments.

The methodology, in particular the operationalization and exploration part, is still under development, as is this particular case study. We are currently finalizing a *predictive processing toolbox* of algorithms for the various sub-processes in predictive processing (to be publicly available) and initiating a number of research projects that build on this work. In particular we aim to explore the effects of various mechanisms for prediction error minimization (model revision, belief revision, passive intervention and active intervention) and the effects of changing the state space granularity in making predictions, allowing for more informative predictions with the drawback of potentially increasing prediction errors. Both aspects are identified as key open theoretical problems within the predictive processing account [9, 11].

References

[1] A. Clark. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3):181–204, 2013.

[2] Open Science Collaboration. Estimating the reproducibility of psychological science. *Science*, 349(6251), 2015.

[3] K.J. Friston. The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010.

[4] L. Heil, S. van Pelt, J. Kwisthout, I. van Rooij, and Harold Bekkering. Higher-level processes in the formation and application of associations during action understanding. *Behavioral and Brain Sciences*, 2(37):202–203, 2014.

[5] J. Hohwy. *The Predictive Mind*. Oxford University Press, 2013.

[6] E. Kayhan, L. Heil, J. Kwisthout, I. van Rooij, S. Hunnius, and H. Bekkering. What's next?: Young children integrate prior probability information with new observations to predict others' actions. Poster presentation at *SRCD'15*, 2015.

[7] J. M. Kilner, K. J. Friston, and C. D. Frith. Predictive coding: An account of the mirror neuron system. *Cognitive Process*, 8:159–166, 2007.

[8] J. Kwisthout. Minimizing relative entropy in hierarchical predictive coding. In L.C. van der Gaag and A.J. Feelders, editors, *Proceedings of PGM'14*, LNCS 8754, pages 254–270, 2014.

[9] J. Kwisthout, H. Bekkering, and I. van Rooij. To be precise, the details don't matter: On predictive processing, precision, and level of detail of predictions. *Brain and Cognition*, under review.

[10] J. Kwisthout, M. Otworowska, H. Bekkering, and I. van Rooij. Leaving Andy Clark's safe shores: Scaling predictive processing to higher cognition. Poster presentation at *CogSci'14*, 2014.

[11] J. Kwisthout and I. van Rooij. Free energy minimization and information gain: The devil is in the details. *Cognitive Neuroscience*, 6(4), 2015.

[12] A. Newell, J. C. Shaw, and H. A. Simon. Report on a general problem-solving program. In *Proceedings of the ICIP*, pages 256–264. Paris: UNESCO House., 1959.

[13] J.X. O'Reilly, U. Schüffelgen, S. Cuell, T.E.J. Behrens, R.B. Mars, and M.F.S. Rushworth. Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences USA*, 110:E3660–E3669, 2013.

[14] M. Otworowska, J. Kwisthout, and I. van Rooij. Counter-factual mathematics of counterfactual predictive models. *Frontiers in Consciousness Research*, 5:801, 2014.

[15] J. Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge: MIT Press, 2000.

[16] D. Rasmussen and C. Eliasmith. God, the devil, and the details: Fleshing out the predictive processing framework. *Behavioral and Brain Sciences*, 36(3):223–224, 2013.

[17] S. van Pelt, L. Heil, S. Ondobaka, J. Kwisthout, I. van Rooij, and H. Bekkering. Neuromagnetic correlates of action probabilities at different hierarchical levels. Poster presentation at *Biomag'14*, 2014.

[18] Springer Verlag. *Language Grounding in Robots*. 2012.

[19] B. H. Webb. Can robots make good models of biological behavior? *Behavioral and Brain Sciences*, 24:1033–1094, 2001.

[20] L. Zaadnoordijk, S. Hunnius, M. Mayer, J. Kwisthout, and I. van Rooij. The developing sense of agency: An interdisciplinary challenge. Poster presentation at *IDCL/EpiRob'15*, 2015.

[21] L. Zaadnoordijk, S. Hunnius, M. Mayer, J. Kwisthout, and I. van Rooij. What senses of agency can infants have? Poster presentation at *CogSci'15*, 2015.