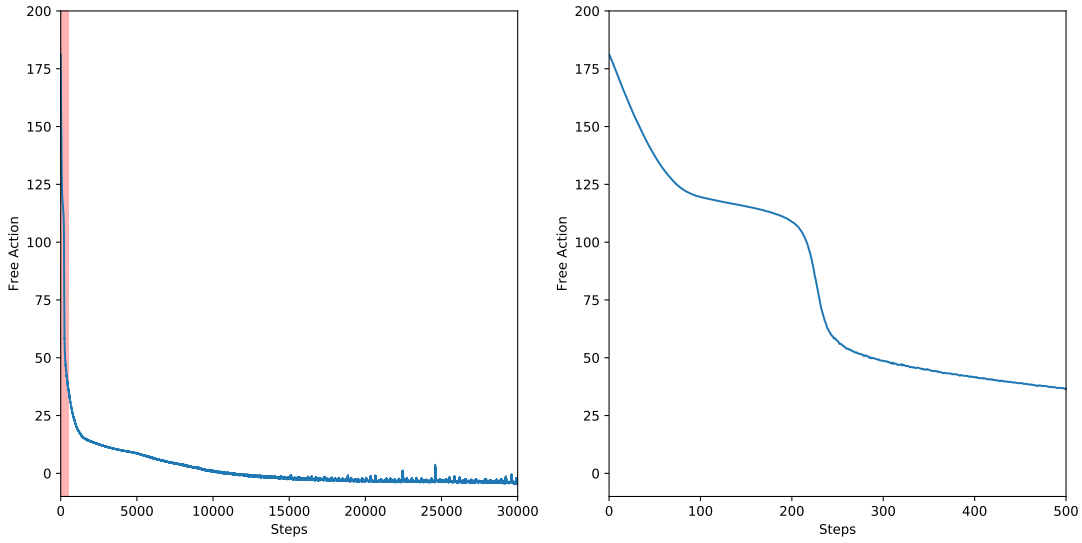
SUPPLEMENTARY MATERIAL TO DEEP ACTIVE INFERENCE

KAI UELTZHÖFFER

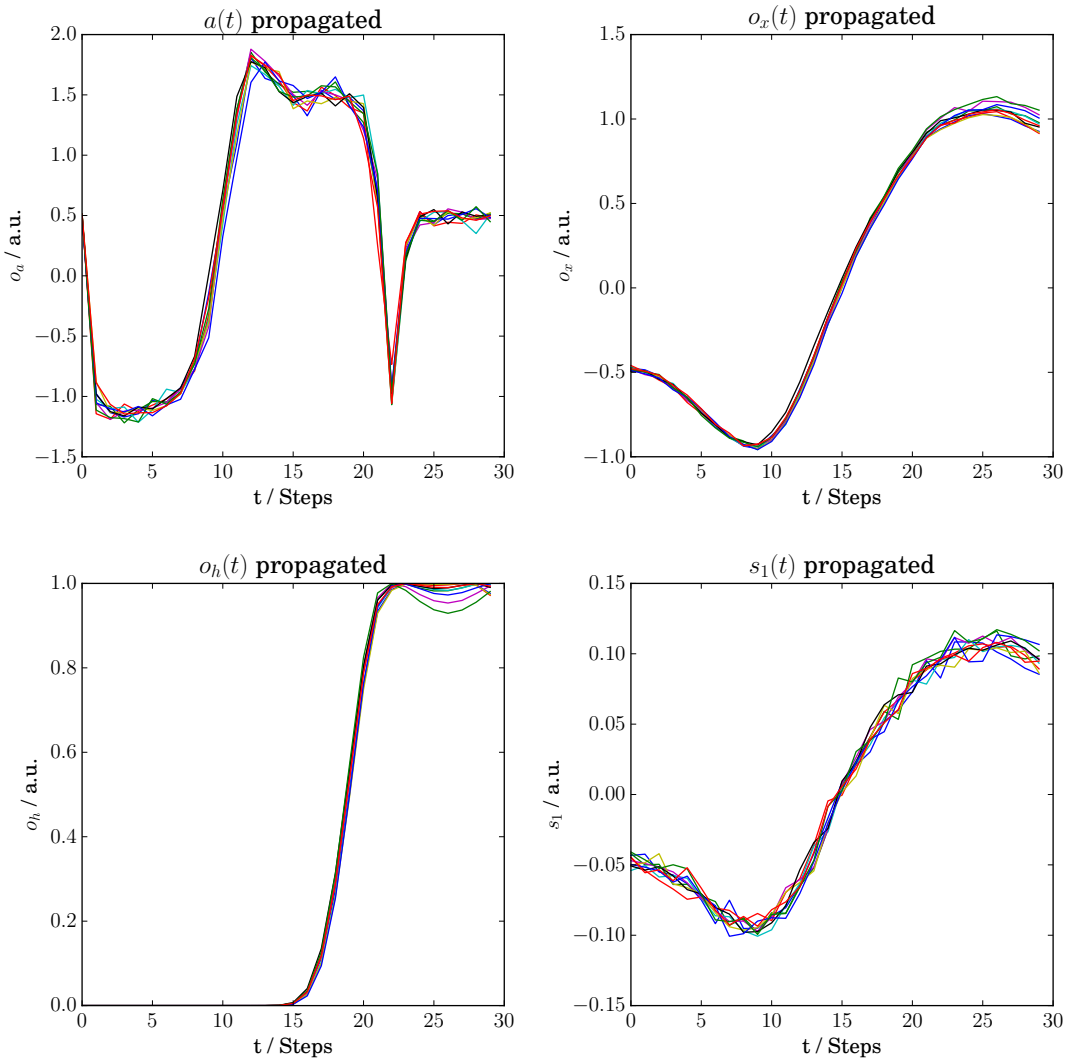
Date: September 18, 2018.

1. Performance without an explicit proprioceptive channel

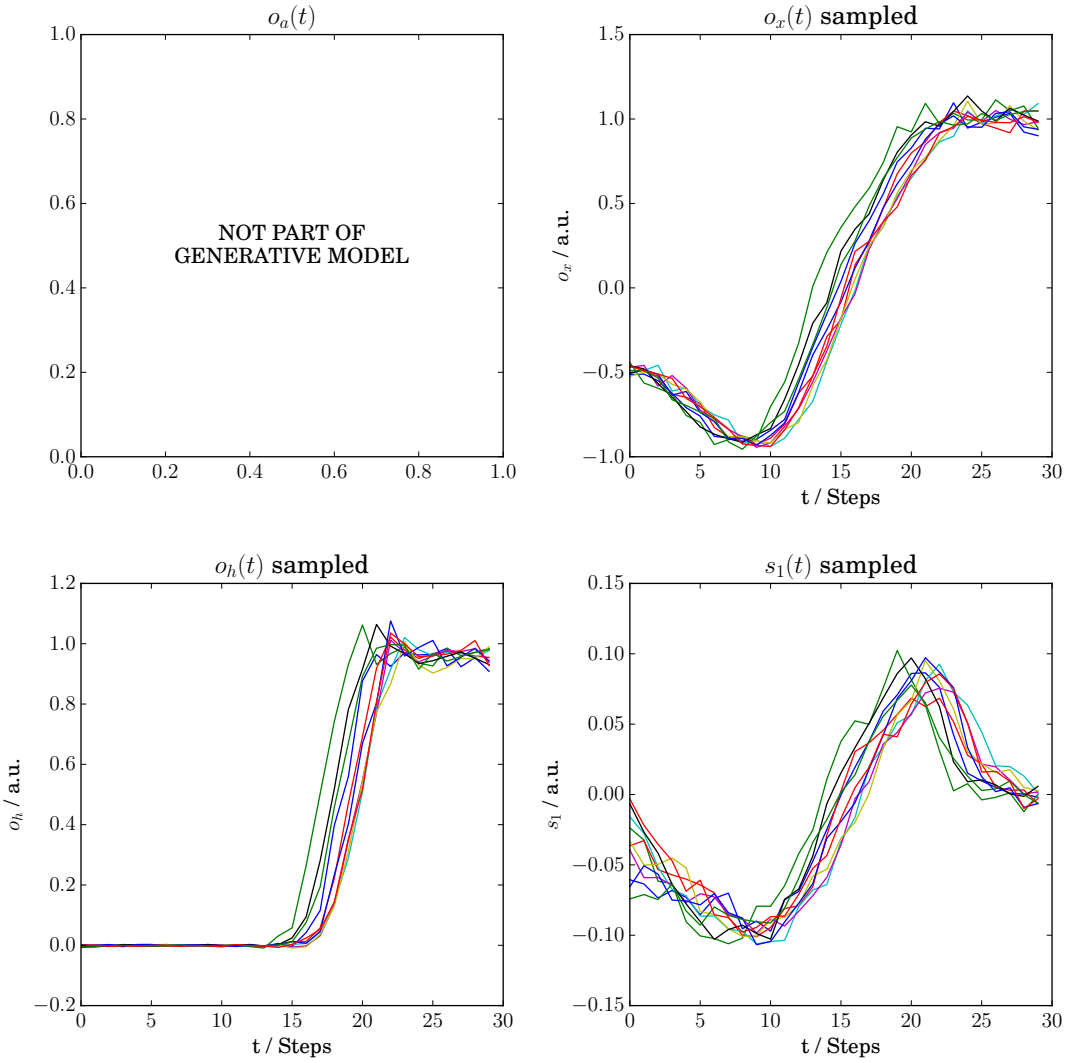
Even without direct feedback on its actions, in terms of a proprioceptive sensory channel o_a, our agent is able to successfully learn the goal instilled in terms of its prior expectations, while simultaneously building a generative model of its exteroceptive sensations. The convergence of the free energy bound is shown in supplementary figure 1. The true interaction of the agent with its environment after 30,000 training steps is shown in figure 2, and its generative model of the world in figure 3. The full code can be accessed at `http://www.github.com/kaiu85/deepAI_paper`.



Supplementary Figure 1. Convergence of an active inference agent, which does not possess a proprioceptive sensory channel, i.e. which gets no direct feedback on his actions. The area shaded in red in the left plot was enlarged in the right plot.



SUPPLEMENTARY FIGURE 2. Performance of an agent without a proprioceptive sensory channel o_a after 30,000 training steps, using the mean parameters of the population density. The agent has acquired a very efficient strategy to reach its goal position: it swings a bit to the left and then directly swings up to its goal position $x = 1.0$. Shown are the agent's action a (upper left), its sense of position o_x (upper right), its nonlinearly transformed sensory channel o_h and its "homeostatic" hidden state s_1.

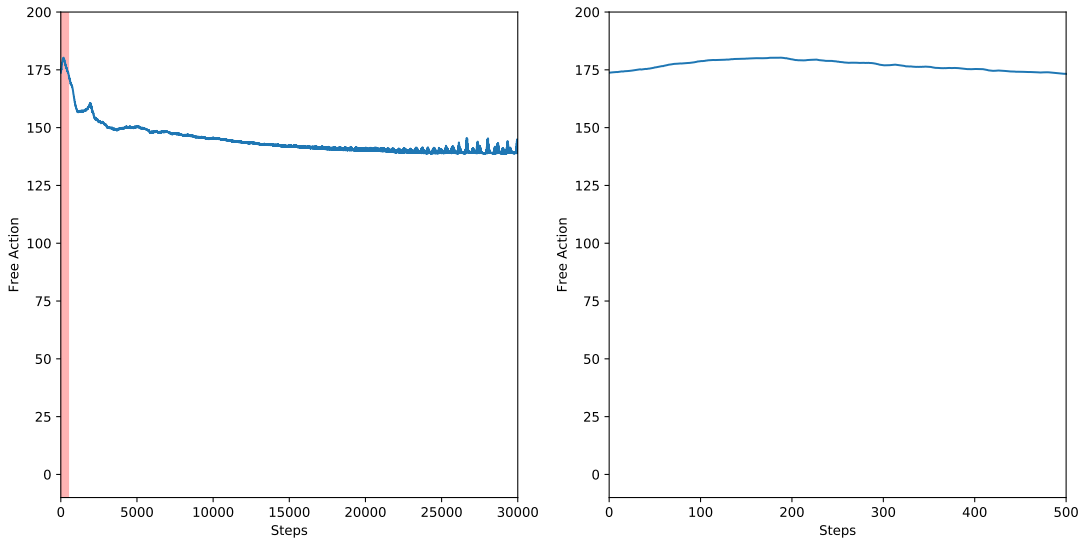


SUPPLEMENTARY FIGURE 3. Sample from the generative model acquired by an agent which does not possess a proprioceptive sensory channel o_a after 30,000 training steps. Shown are the agent's sense of position o_x (upper right), its nonlinearly transformed sensory channel o_h and its "homeostatic" hidden state s_1. Note the very nice correspondence to its actual trajectory, when interacting with the world, as shown in supplementary figure 2

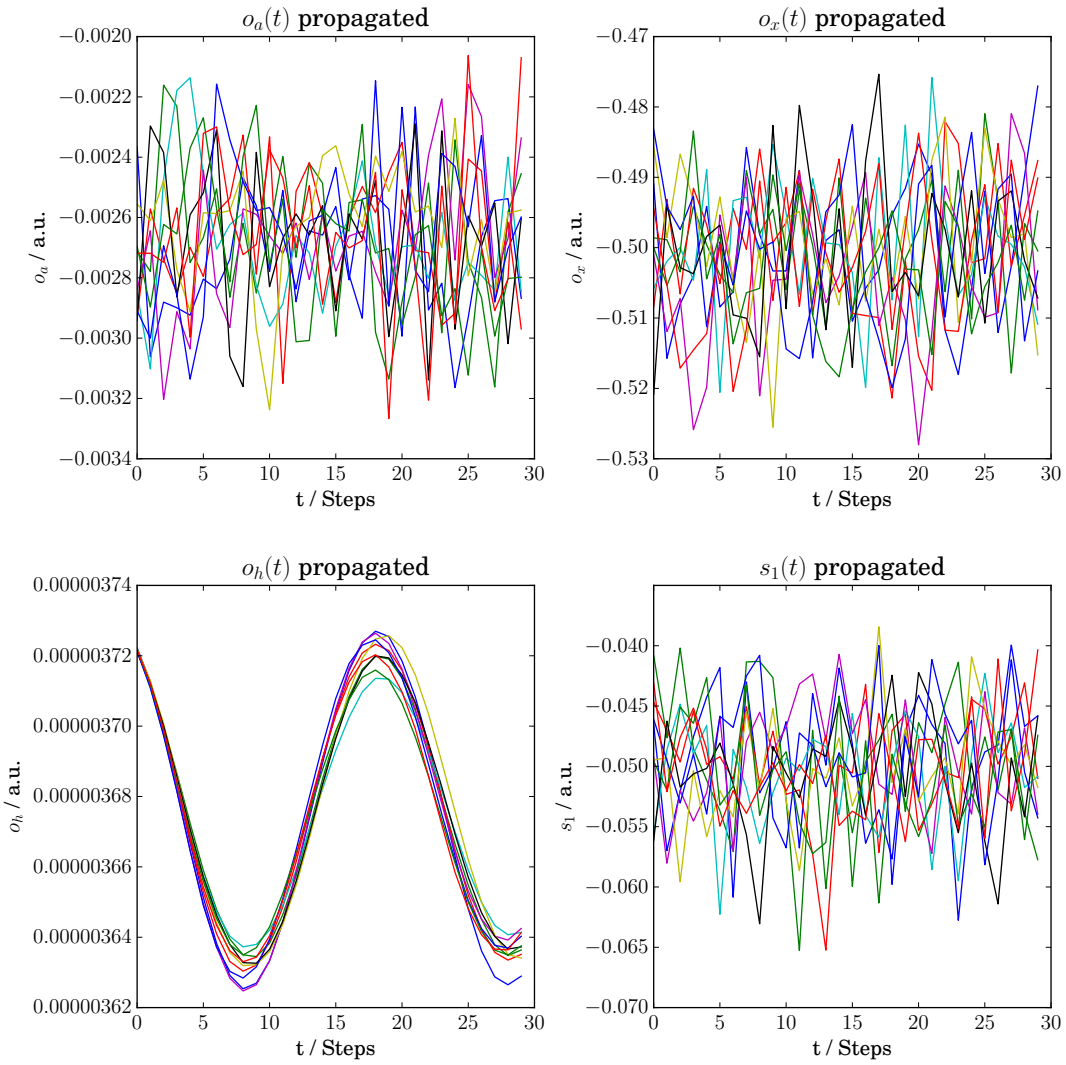
2. Performance with purely proprioceptive action

If action is used only to directly suppress proprioceptive surprise, the convergence of the learning process is severely impaired, as shown in supplementary figure 4. Here we optimise an agent whose structure, parameters and objective function are identical to the Deep Active Inference agent in the main text. However, the updates on the parameters of the agent's action function only depend on the gradient of the expectation value over the population density of the sensory surprise in the proprioceptive channel $< -\ln p_\theta(o_{a,t}|\mathbf{s}_t) >_{q(\mathbf{s}_t|\mathbf{s}_{t-1},o_{x,t},o_{h,t},o_{a,t})}$ with respect to the parameters of the action function. This corresponds to an agent which neglects the direct changes in other sensory modalities due to its actions. One example might be the complex, nonlinear changes in the visual input to the retina, which arise even from small eye movements. The full code can be accessed at `http://www.github.com/kaiu85/deepAI_paper`.

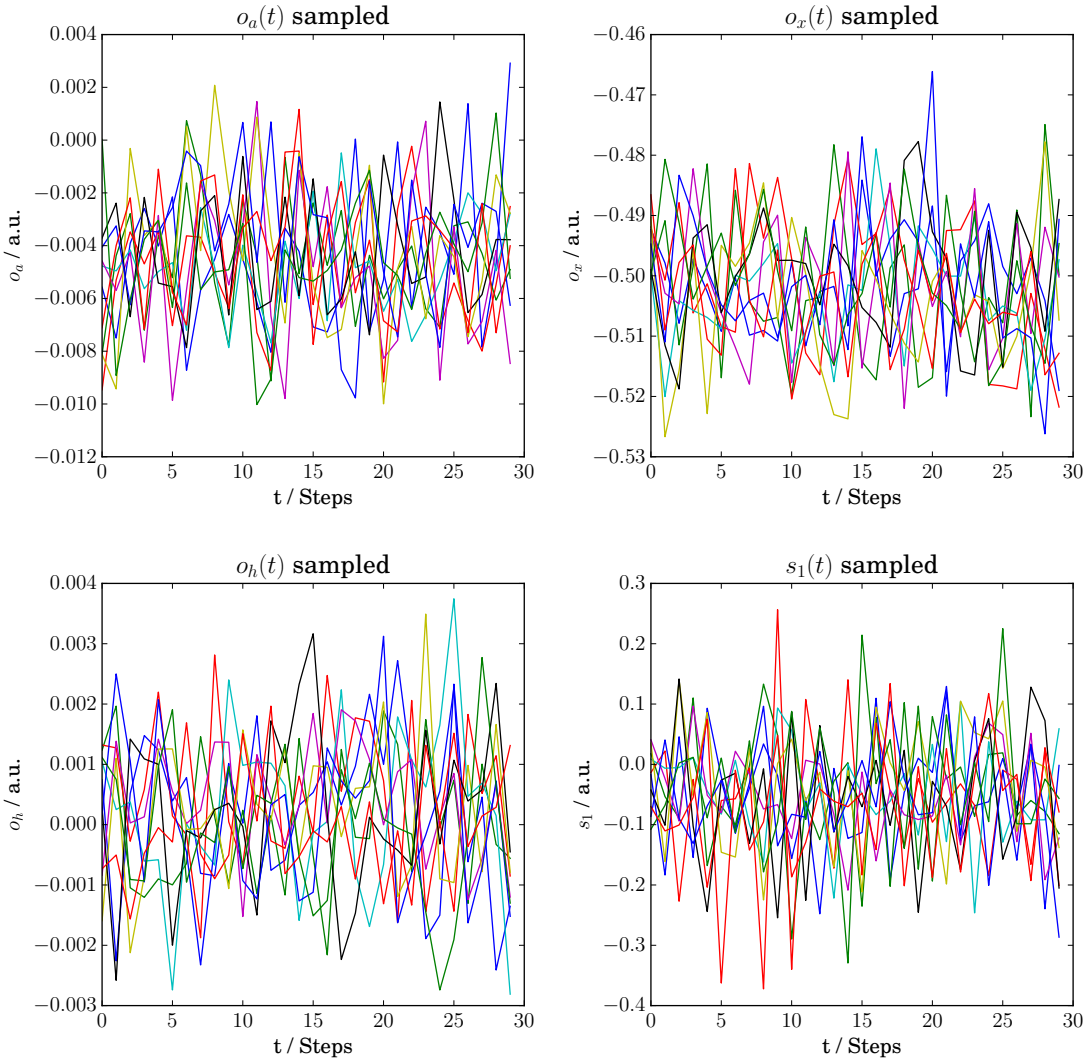
Comparing supplementary figure 4 to figure 3 in the main text or to supplementary figure 1, which shows the convergence of an active inference agent lacking any proprioceptive input, it is obvious that this reduction prevents the agent from successfully achieving its goals and learning about its environment. This is also seen in the behavior of such an agent after 30,000 training steps, shown in supplementary figure 5, and its - nonexistent - generative model of the world shown in supplementary figure 6.



Supplementary Figure 4. Convergence (or rather the lack of it) of an active inference agent, which uses action only to directly suppress its proprioceptive surprise. The area shaded in red in the left plot was enlarged in the right plot.



Supplementary Figure 5. Performance of an agent which uses action only to directly suppress its proprioceptive surprise after 30,000 training steps, using the mean parameters of the population density. The agent is stuck at its initial position and shows no clear behavioral strategy. Shown are the agent's proprioceptive channel o_a (upper left), its sense of position o_x (upper right), its nonlinearly transformed sensory channel o_h and its "homeostatic" hidden state s_1.



SUPPLEMENTARY FIGURE 6. Sample from the generative model acquired by an agent which uses action only to directly suppress its proprioceptive surprise after 30,000 training steps. Shown are the agent's proprioceptive channel o_a (upper left), its sense of position o_x (upper right), its nonlinearly transformed sensory channel o_h and its "homeostatic" hidden state s_1. Note the lacking correspondence to its actual trajectory (e.g. in o_h or s_1), when interacting with the world, as shown in supplementary figure 5