Derivation of 'optimal weighting' probability (pO),
Hunt et al., Nat Neurosci 2014

Considering a particular option *i* in the three-option choice task, let
H_1 be the hypothesis that *i* is rewarded;
H_0 be the hypothesis that *i* is unrewarded.

Let s_i denote the stimulus at option *i*, and a_i denote the action at option *i.*
$p(H_1|s_i)$ is the learnt stimulus-reward probability - denoted pS_i in the paper.
$p(H_1|a_i)$ is the learnt action-reward probability - denoted pA_i in the paper.

We want to determine $p(H_1|s_i,a_i)$ - denoted pO_i in the paper.

We assume that the prior probability that the option is rewarded, $p(H_1)$ is 0.5.

Also, as H_1 and H_0 are the only two possible outcomes,
$p(s_i|H_0) = (1\text{-}p(s_i|H_1))$, and $p(a_i|H_0) = (1\text{-}p(a_i|H_1))$.

Similarly, $p(H_0) = 1\text{-}p(H_1) = 0.5$.

First we use Bayes' theorem to show that $p(s_i|H_1) = p(H_1|s_i)$:

$$p(H_1|s_i) =$$

$$\frac{p(s_i|H_1)p(H_1)}{p(s_i|H_1)p(H_1) + p(s_i|H_0)p(H_0)} =$$

$$\frac{p(s_i|H_1) * 0.5}{p(s_i|H_1) * 0.5 + (1 - p(s_i|H_1)) * 0.5} =$$

$$\frac{p(s_i|H_1)}{p(s_i|H_1) + \left(1 - p(s_i|H_1)\right)} =$$

$$p(s_i|H_1)$$

Similarly, $p(a_i|H_1) = p(H_1|a_i)$.

Second, as actions and stimuli are conditionally independent of each other, we know that $p(s_i,a_i|H_1) = p(s_i|H_1)*p(a_i|H_1)$. Similarly, $p(s_i,a_i|H_0) = p(s_i|H_0)*p(a_i|H_0)$.

From Bayes' theorem, we can therefore calculate $p(H_1|s_i,r_i)$ as follows:

$$p(H_1|s_i,a_i) = \frac{p(s_i,a_i|H_1)p(H_1)}{p(s_i,a_i|H_1)p(H_1) + p(s_i,a_i|H_0)p(H_0)}$$

$$= \frac{p(s_i|H_1)p(a_i|H_1) * 0.5}{p(s_i|H_1)p(a_i|H_1) * 0.5 + \ p(s_i|H_0)p(a_i|H_0) * 0.5}$$

$$= \frac{p(s_i|H_1)p(a_i|H_1)}{p(s_i|H_1)p(a_i|H_1) + (1 - p(s_i|H_1))(1 - p(a_i|H_1))}$$

$$= \frac{p(H_1|s_i)p(H_1|a_i)}{p(H_1|s_i)p(H_1|a_i) + (1 - p(H_1|s_i))(1 - p(H_1|a_i))}$$

Or, using the notation in the paper:

$$pO_i = \frac{pS_i.pA_i}{pS_i.pA_i + (1 - pS_i).(1 - pA_i)}$$