Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

Bayesian cue integration as a developmental outcome of reward mediated learning

Thomas H Weisswange, Constantin Rothkopf, Tobias Rodemann, Jochen Triesch

2011

Preprint:

This is an accepted article published in PLoS ONE. The final authenticated version is available online at: https://doi.org/[DOI not available]

OPEN ACCESS Freely available online

PLoS one

Bayesian Cue Integration as a Developmental Outcome of Reward Mediated Learning

Thomas H. Weisswange[1]*, Constantin A. Rothkopf[1], Tobias Rodemann[2], Jochen Triesch[1]

1 Frankfurt Institute for Advanced Studies, Frankfurt, Germany, **2** Honda Research Institute Europe GmbH, Offenbach, Germany

Abstract

Average human behavior in cue combination tasks is well predicted by Bayesian inference models. As this capability is acquired over developmental timescales, the question arises, how it is learned. Here we investigated whether reward dependent learning, that is well established at the computational, behavioral, and neuronal levels, could contribute to this development. It is shown that a model free reinforcement learning algorithm can indeed learn to do cue integration, i.e. weight uncertain cues according to their respective reliabilities and even do so if reliabilities are changing. We also consider the case of causal inference where multimodal signals can originate from one or multiple separate objects and should not always be integrated. In this case, the learner is shown to develop a behavior that is closest to Bayesian model averaging. We conclude that reward mediated learning could be a driving force for the development of cue integration and causal inference.

Citation: Weisswange TH, Rothkopf CA, Rodemann T, Triesch J (2011) Bayesian Cue Integration as a Developmental Outcome of Reward Mediated Learning. PLoS ONE 6(7): e21575. doi:10.1371/journal.pone.0021575

Editor: Eleni Vasilaki, University of Sheffield, United Kingdom

Received February 10, 2011; **Accepted** June 3, 2011; **Published** July 5, 2011

Copyright: © 2011 Weisswange et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: Authors J.T. and C.R. were supported by the EC MEXT project PLICON; J.T., C.R., and T.W. by the German Federal Ministry of Education and Research (BMBF) within the "Bernstein Focus: Neurotechnology" through research grant 01GQ0840. These funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Author T.W. gratefully acknowledges the financial support from Honda Research Institute Europe GmbH. Author T.R. is an employee of Honda Research Institute Europe GmbH and played a role in design and writing of the paper.

Competing Interests: Author T.R. is an employee of the Honda Research Institute Europe GmbH. He does not endorse any Honda products or opinions in his research or in reports of his research results. There are no patents, products in development or marketed products to declare. All authors adhere to all the PLoS ONE policies on sharing data and materials, as detailed online in the guide for authors.

* E-mail: weisswange@fias.uni-frankfurt.de

Introduction

Empirical studies have provided convincing evidence that humans combine sensory signals into percepts so as to reduce the uncertainty about their causes. Such studies, as reviewed in [1–6], have commonly used the cue integration paradigm, in which human observers are asked to infer a certain quantity based on the observations of a bi– or multisensory signal. These experiments may use stimuli across modalities such as in the judgment of the position of an object based on visual and auditory cues [7,8] or object size given visual and haptic cues [9]. Similarly, experiments have considered cues within the same modality as in inferring surface slant from stereo and texture cues [10] or depth from texture and motion cues [11]. The overwhelming majority of these studies has shown that humans combine these cues by weighting them according to their relative reliabilities.

This empirical behavior has been modeled as inference of the most likely cause of the observation given the sensory cues and prior knowledge using the Bayesian framework. Bayesian inference represents the uncertainty about parameters in the inferential task explicitly as probability distributions. The variance of the distribution resulting from optimally combining the different sources of information is smaller than that of the individual sensory distributions which reflects the reduction in uncertainty.

The Bayesian inference framework has recently been extended to cases in which the observed sensory signals are caused by one of two different scene configurations [12,13]. As an example, consider the case in which a visual and an auditory cue are sensed. Whether these cues should be integrated or not may depend on the assumptions about the causes for these signals. If two separate objects caused them, they should not be combined, whereas if the two signals are likely to be caused by a single object in the scene, they should be integrated. By representing the uncertainty about the two possible scene configurations, the Bayesian framework can be used to compute a posterior probability distribution over the two scene layouts, a process that has been termed causal inference [13,14]. Instead of inferring the scene layout, the task may be to compute a posterior probability of the positions of the signal source by weighting the likelihoods of the two scene layouts. While model selection is the optimal strategy for the former task, the latter task favors Bayesian model averaging. In most experimental setups however, it is difficult to distinguish which of these strategies best matches human behaviour or whether humans use a fixed single strategy at all [15].

Despite the aforementioned successes in applying the Bayesian framework to sensory perception, a wide variety of questions remain open at the computational, algorithmic, and implementation levels, see e.g. [6,16]. The most central issue is how cue integration and causal inference are learned on developmental timescales [17–20]. A recent empirical study by Gori and colleagues [17] showed that children under 10 years of age did not integrate cues by taking their uncertainties into account. Instead, these children were shown to almost exclusively use haptic cues in a size discrimination task and visual cues in an orientation discrimination task independently of their reliability. This suggests that the abilities for cue integration and causal inference are

acquired through development. This result is particularly challenging for some current theoretical models at the implementational level suggesting that cue integration is mediated simply by the Poisson-like trial to trial variability in neuronal populations [21,22]. These so called probabilistic population codes were shown to be able to integrate probability distributions, represented by probabilistic activity within specific neuronal populations, using simple biologically plausible computations. However, it is unclear how such a mechanism could be learned over developmental timescales. It is an open question then, why infants and children do not integrate cues optimally and how learning could proceed [16].

Further evidence for the role of learning in the development of sensory integration and causal inference comes from a study by Putzar et al. [23] in which the authors show that early deprivation of one modality during the first month of life impaired multisensory integration including this sense even after complete recovery (see also [24]). This matches neurophysiological findings in cats and monkeys suggesting a critical period of high plasticity and large changes in receptive fields of multisensory neurons during early development [25,26]. Neurons in superior colliculus of newborn kittens and monkeys show little or no multisensory responses, but the number of multisensory neurons grows and their tuning gets sharper with age.

There is also initial evidence, that the mechanisms involved in causal inference are not fully developed at birth, at least in cats [27]. Cats were raised in an artificial environment, in which auditory and visual signals were always shown at the same time but at differing spatial positions. Subsequent behavioral as well as neurophysiological tests revealed that the animals did not integrate multisensory stimuli from a common location, as seen in animals raised in natural environments, but instead integrated only signals with the distinct spatial separation present in the artificial environment. For similar results in owls see [28].

Based on these results, the present study asks, whether cue integration and causal inference in sensory perception could develop mediated by reward dependent learning. There are ample data demonstrating reward dependent learning related to orienting movements [29–31]. Furthermore, there has been considerable work relating this type of learning to theoretical models of reinforcement learning (RL), see e.g. [32]. Various studies were able to localize areas in the human and monkey brain potentially implementing RL mechanisms by looking for correlations between RL model variables and single cell [32,33] or BOLD activities [34–40]. Those studies show the relevance of RL for learning in many different tasks and environments both on the behavioral as well as the computational level. Thus it is interesting to also consider it as one potential driving force, on the computational level, for the development of cue integration and causal inference.

Results

We use a multimodal localization task similar to the one used by Neil and colleagues [20] and Körding et al [13] (see Fig. 1 for a schematic depiction). The learner obtains noisy visual and auditory signals and carries out horizontal orienting movements, obtaining a varying amount of reward dependending on the accuracy of the movement (see Methods). We interpret the reward as an intrinsic signal for bringing a relevant stimulus into the center of attention.

The agent learns to solve this task based only on its sensory inputs, orienting actions, and observed rewards. To this end, it learns to predict how much reward to expect when performing each action in a given situation. The learner represents its reward estimates for particular state and action pairs as so called Q-values [41]. Support for the representation of such variables in the human and monkey brain comes from several studies [33,42]. In our case this Q-function is approximated by a neural network (see Methods). Based on these reward expectations, the agent will probabilistically pick an action using a softmax function, which also has been shown to match human action selection for some tasks [43,44]. The reward prediction of the winning action will be adapted depending on the difference between predicted and obtained reward by changing all synaptic weights via a gradient descent learning algorithm (see Methods).

In the following we will test this model on cue integration and causal inference tasks and compare it to human behavior and four different Bayesian models.

Cue Integration

We start with a simple cue integration paradigm, where noisy auditory and visual signals from a common source have to be

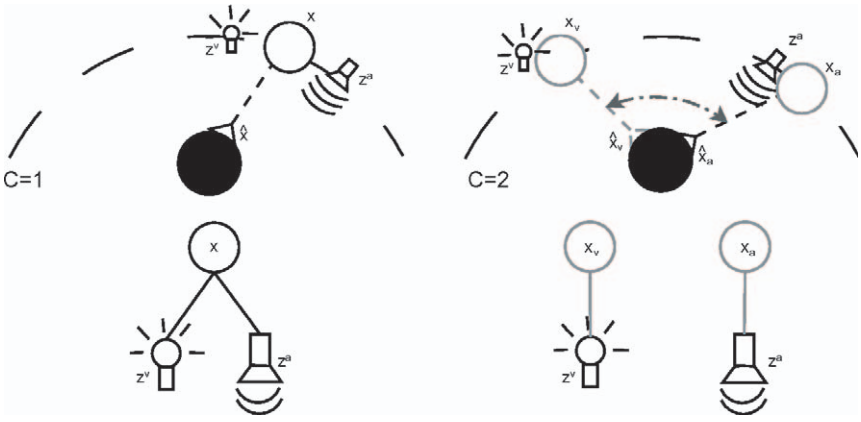

Figure 1. Scene layout of orienting tasks and generative models. Top: Sketch of the orienting task used in this study. The learner receives an auditory (z^a) and a visual (z^v) signal, which are probabilistically related to the true position x. The task is to orient towards this true position. Left: A case where the visual and auditory signals have a common cause. Right: The signals originate from different locations. Bottom: The generative models for the task. The noisy sensory signals (z^a and z^v) are either generated by a single ($C=1$) or two independent causes ($C=2$) having different spatial positions.
doi:10.1371/journal.pone.0021575.g001

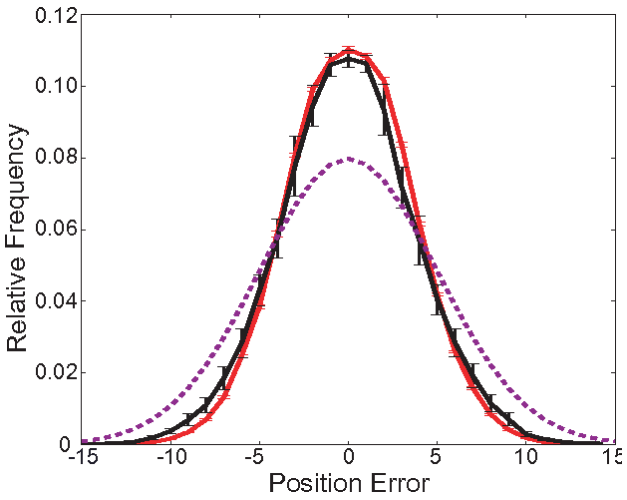

Figure 2. Distribution of position estimation errors. The distribution of errors over 100,000 orienting actions carried out by the RL model after 10 runs with each 100,000 training steps (black), compared with Bayesian optimal integration (red) and the best single cue predictions (dashed) for a single audio-visual object. Errorbars show standart deviation of 10 runs. ($\sigma_a^2 = 5$, $\sigma_v^2 = 5$).
doi:10.1371/journal.pone.0021575.g002

combined. If the noise of the two cues is independent, the variance of the error produced by optimally integrating the two stimuli is always smaller or equal to the error variance resulting from using either cue alone. Figure 2 shows the distribution of errors the RL based model produces after training. This result matches well with the predictions of the optimal Bayesian model for this situation.

To compare the model with human behavior, we test the fully trained model on a two-alternative forced choice (2-AFC) task. This task allows us to test the behavior of the learner for changes in relative reliabilities between the cues. The setup is similar to the one used by Ernst and Banks [9], where human subjects were asked to perform a 2-AFC visuo-haptic size discrimination task. Ernst and Banks could show that in this task the point of subjective equality (PSE) of adults is well predicted by Bayesian cue integration and shifts when additional visual noise is introduced.

The first input to the agent is the size of a standard bar with constant position, the second is the size of a probe which varies and is to be estimated as 'left' or 'right' of the standard (respectively 'taller' or 'smaller' in [9]). Both stimuli are bimodal, but for the probe the cues are always consistent, whereas for the standard they are set to be in conflict with each other. Figure 3A shows the proportion of 'right' estimates for all possible positions of the probe based on the decisions taken by the reinforcement learner after training as psychometric curves. Each curve represents training and testing with a different visual noise variance. We can compare it with the data from of Ernst and Banks [9] (Figure 3B) which is reproduced in our Figure 3B. It shows the equivalent data for the average of four human subjects. Both plots show a similar pattern in that the psychometric curves get steeper and the PSE moves more towards the visual stimulus position for decreasing visual noise levels. Figure S2 compares the PSEs of the RL model (crosses) with that of the optimal Bayesian observer (circles) for different visual reliabilities. It can be seen that they match quite well, as was true for the human subjects in [9] (Fig 3c in their paper). Note that there is variability in both the PSEs of the learner and the Bayesian observer due to the limited number of test stimuli.

Causal Inference

In the following tasks we will add a second layer of complexity to the task by randomly presenting trials that were generated by different scene layouts, i.e. under either the common or the separate cue condition. We will compare our learned model with four Bayesian observers. One observer always integrates the information from the two stimuli (we will call that one AI). A second always acts as if both stimuli originate from different objects and discards information from the less reliable modality ("Never Integrating" – NI). A third, more advanced, observer computes the probability of one vs. two objects in each trial and uses the optimal action for the more probable model ("Model Selection" – MS). The fully optimal fourth observer though makes use of all information available by selecting an action under the weighted evidence for each generative model, with the weights proportional to the respective probabilities ("Model Averaging" – MA). All Bayesian observers, contrary to the RL model, have explicit knowledge of position priors, sensory noise distributions

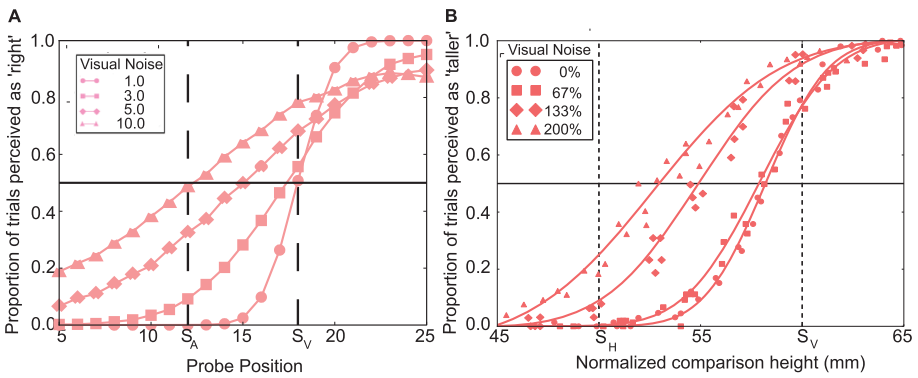

Figure 3. Psychometric curves for 2-AFC task in comparison to human psychophysics. A: Psychometric curves for the proportion of 'right'-actions in an audio-visual 2-AFC position estimation task. The input positions of the standard were mismatched, with the auditory signal positioned at 12 and the visual signal at 18. Probe inputs were matched and tested 1,000 times at each position (0–30). The point at which the curves cross the black vertical line is the PSE. The curves differ in the variance of the visual noise (see legend), auditory noise was kept constant with $\sigma_a^2 = 5$. B: Plot using data from a psychophysical experiment by Ernst and Banks [9]. They used a visuo-haptic 2-AFC size discrimination task and count the proportion of 'taller'-actions. The standard inputs were mismatched (haptic at 50, visual at 60), probe inputs were matched and varied between 45 and 65. The visual reliability was varied by adding external noise to the display.
doi:10.1371/journal.pone.0021575.g003

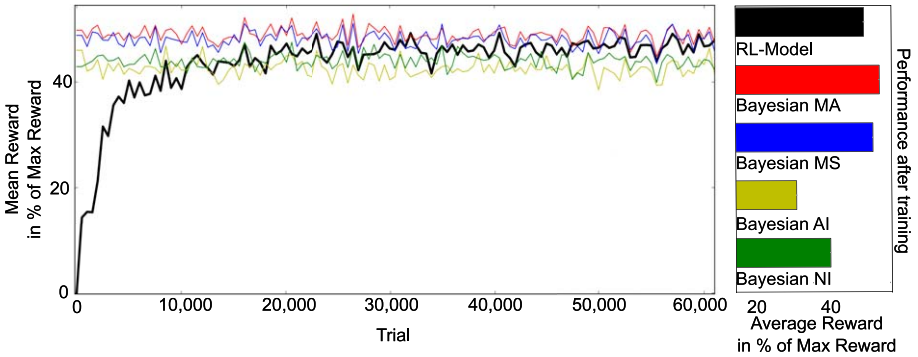

Figure 4. Performance of the RL model and Bayesian observers for a single output. Reward obtained by the learner when choosing the action with highest predicted reward (black) compared to the different Bayesian observers. Signals can originate from one or two objects. Left: Change of performance during learning. Each data point is the sliding average of 1000 trials. Right: Barplot of the mean reward over 100,000 trials after learning. Standard error of the means is smaller than 0.5% for all bars. ($\sigma_a^2=3$, $\sigma_v^2=2$).
doi:10.1371/journal.pone.0021575.g004

and the reward rule. The mathematical formulation of these decision rules as well as the reward expectations of the observer models can be found in the Methods section.

To show the learning process of the RL learner, we can look at the development of the potential reward received with a greedy policy (always selecting the action which predicts the highest reward). In Figure 4 one can see that the average reward earned by the learner increases until it reaches a level similar to what the MA and MS models show (see also Table 1 1A). Comparing it with the simpler instances of Bayesian observers, the learner is clearly better than AI and NI, that is, it implicitly incorporates the existence of two different conditions. But it is hard to tell apart the Bayesian MA and MS observers. Both are similar to the agent's performance in this task.

A different way of assessing the behavior of the RL agent is to directly consider the expected total discounted reward obtainable for a particular state-action pair, the Q-values. Figure 5 shows subsections of two learned Q-value function approximations for all inputs, a given action nd two different reliability ratios. The highest reward is expected if both input signals are close together, resulting in a high probability for a single cause, and close to the target of the given action, resulting in a high probability for the action being correct (Fig. 5 center of both plots). Importantly, if the target of the given action can not possibly be a result of a weighted average of the input positions – because the cues favour both a higher or both a lower position, this action predicts little reward (Asterisks in Fig. 5). For this reason the plots show an asymmetric reward landscape. The slant of the area of highest reward (dark red) depends on the relative reliability of the two cues, as can be seen when comparing A and B in Fig. 5. The left plot is a result of inputs with a higher reliability in the visual modality, therefore the area of highest reward lies more along the visual axis, whereas in the right plot with equal reliabilities for both cues it lies along the diagonal exactly between the auditory and visual axis. The width of this area, as well as the maximum predicted reward, is determined by the absolute values of the reliabilities (narrower and darker red in the left plot due to higher visual reliability). A smaller reward can be expected if the cues are far apart – resulting in a high probability for two causes, but one of them is close to the action target – resulting in a high probability for the action to be correct for one of the objects (Middle of each of the four figure boundaries in Fig. 5 – the "arms" of the cross). The height of these expectations depends again on the reliability of the relevant cue.

Table 1. Model performance for different set-ups.

Setup	RL-Model	Bayesian MA	Bayesian MS	Bayesian MI	Bayesian NI
1A	46.62%	47.9%	47.04%	37.87%	41.57%
1B	46.32%	47.9%	47.04%	37.87%	41.57%
1C	45.86%	47.06%	46.18%	37.28%	40.88%
2A	50.51%	51.81%	51.08%	41.63%	47.71%
2B	36.26%	37.37%	36.53%	33.87%	34.28%
2C	48.74%	50.52%	49.70%	40.32%	45.67%

Average fraction of maximum reward received in 100,000 steps after learning ($\sigma_a^2=3$, $\sigma_v^2=2$) for different variation the task and the model. Results of the different Bayesian observers for comparison.
doi:10.1371/journal.pone.0021575.t001

In the experimental setup from [13] participants were asked to report in each trial both the visual as well as the auditory location of a stimulus. To mimic this condition, we change our task accordingly and add a second output population to the neural network (see Fig.S1). Each population now represents the actions for one modality (representing the arrays of buttons for the participants in [13]). The rewards and the prediction errors are computed separately (according to (3)). Table 1 2A shows the performance after learning as the sum of both rewards. The effects are similar to the previous orienting task, in that we see a performance thats is similar to the predictions of MA and MS.

Despite changing the task it is still difficult to distinguish these two Bayesian observers (MA and MS) from each other by comparing the collected reward. A better discriminator should be the variance explained by each observer in relation to the total variance of the orienting error of our model (generalized coefficient of determination R^2 [45]). The differences between MA and MS over all inputs are nevertheless still small (Suppl. Fig.S3). Fortunately, since we have the full observer models, we can find the inputs for which the optimal actions differ between MA and MS, and then test the RL model only on those. The R^2 values for these inputs are shown in Fig. 6. We also include an observer which does probability matching (PM) for model selection, proposed to be the strategy used by many human subjects in a recent experiment [15]. It can be seen that the Bayesian observer with model averaging explains the error variances best for both visual and auditory output (grey and black bars). The values for the MS and NI observer are the same,

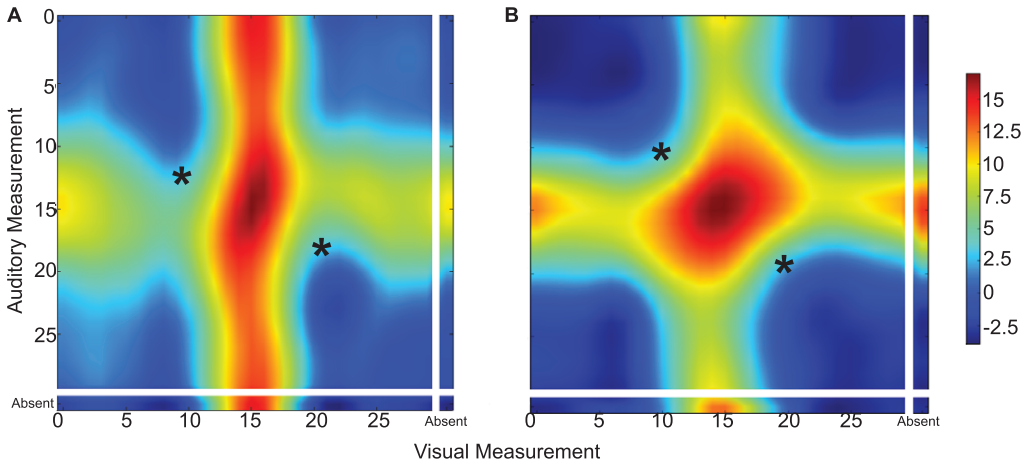

Figure 5. Exemplary subsections of two learned Q-value functions. Expected reward for visual signals (x-axis) and auditory signals (y-axis) for the action of orienting towards the center. Red colors represent high, blue low predicted rewards. Left: Visual cue is more reliable ($\sigma_a^2=3$, $\sigma_v^2=2$); Right: Both cues have the same variance ($\sigma_a^2=\sigma_v^2=3$). For a detailed explanation see main text.
doi:10.1371/journal.pone.0021575.g005

because the selected inputs represent those in which MS decides to act according to the generating model with independent objects.

Complex Uncertainty Structures

While the presented system is certain to accommodate different prior distributions of the relevant scene variable relevant for obtaining rewards (see Table 1 1C for an example with a Gaussian prior for the visual stimulus) because of the maximum total discounted reward guarantees of RL, it is interesting to see whether the system can handle different likelihood landscapes. In many real-life situations, the uncertainty of a cue depends on a number of factors. The following three experiments introduce behaviorally plausible variations in uncertainty structure and investigate how the RL agent can adjust to these.

Spatial Variation in Uncertainty Structure. Visual estimates of spatial location should be more accurate in the fovea than in the periphery of the visual field, given the human acuity falloff (e.g. [46], see also [47] for an example in slant angle space). Figure 7 shows the reward predictions for a set-up that mimics this observation in the task that requires a single action. The variance of the visual noise was low for stimuli in the center and increased with eccentricity, whereas auditory reliability stayed constant (Figure 7 shows results with linear increase of the variance, but similar results are reached with other functions, e.g. logarithmic decay). Training on this adapted task resulted in reward predictions dominated by the visual estimate for actions towards the center (Fig. 7 right) and dominated by audition for the outer periphery (Fig. 7 left). In between these two extremes,

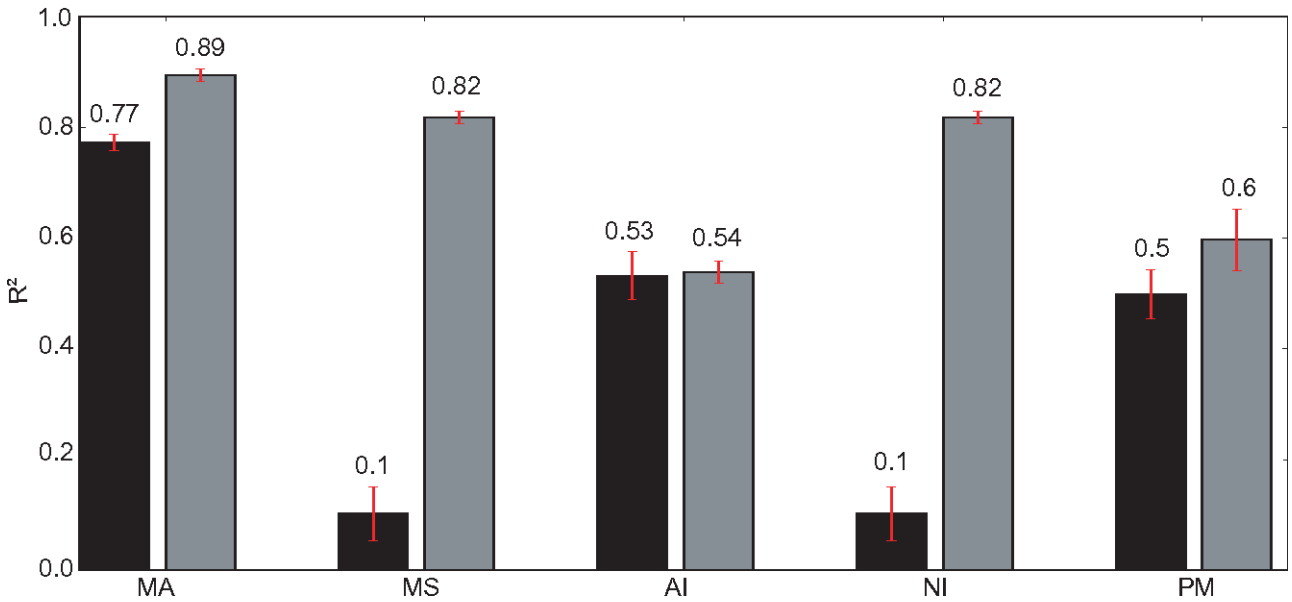

Figure 6. R^2 values of different observers for the RL model with selected inputs. The black and gray bars show the results for the auditory and the visual output for 50,000 trials with inputs that differ in the predicted action between MA and MS. Mean over 10 training sessions with $\sigma_a^2=3$, $\sigma_v^2=2$, errorbars show standard deviation.
doi:10.1371/journal.pone.0021575.g006

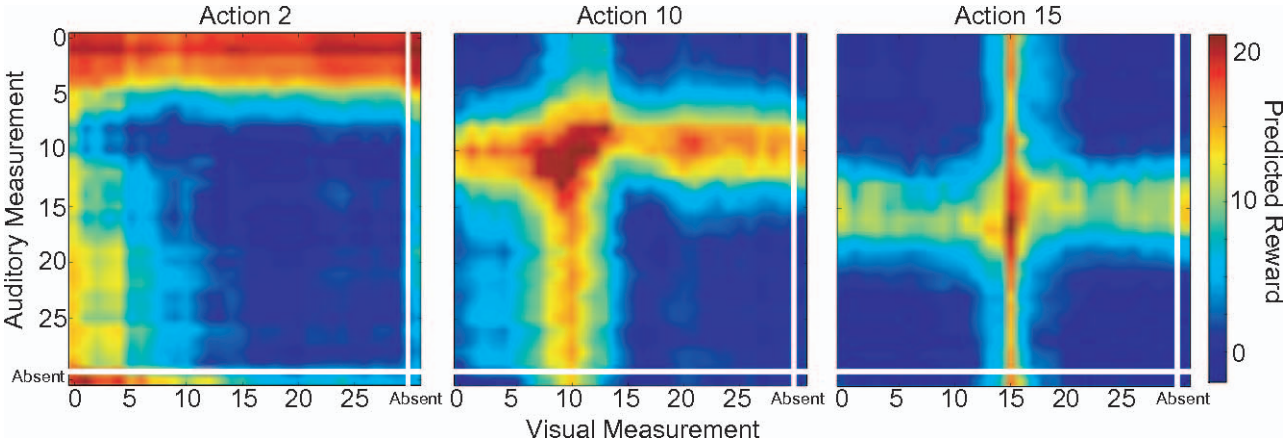

Figure 7. Exemplary subsections of the learned Q-value function for the foveation setup. Axes are the same as Fig. 5, but for three different actions with constant auditory reliability ($\sigma_a^2 = 2$) and space varying visual reliability. L-R: Actions towards a peripheral($\sigma_v^2 = 3.25$), intermediate ($\sigma_v^2 = 2$) and central position ($\sigma_v^2 = 0.25$).
doi:10.1371/journal.pone.0021575.g007

integration of both cues predicted the highest reward. This can also be seen in the distribution of input weights to the hidden layer (Suppl. Fig.S4). The weights from the auditory part of the input layer have similar shapes and width across all positions, whereas the visual weights get narrower towards the central positions. This shows that reward mediated learning results in behavior that varies with context within a single task, which is in accordance with predictions from a Bayesian model that explicitly takes into account context when computing the data likelihood.

Temporal Variation in Uncertainty Structure. In addition to a change in noise variance across space as discussed above, in a natural environment the variance also changes over time. As an example one may consider the change in the optimal weighting of visual compared to auditory cues when stepping out of a dim room into the bright sunlight. Due to higher contrasts and thus smaller uncertainty, visual localization will gain confidence in the latter condition. To simulate such dynamics, we change the reliability of the visual cue at certain timepoints during training (Fig. 8). The network quickly adjusts to a change in visual reliability. The performance after a change point (vertical lines in Fig. 8) quickly becomes similar to the optimal predictions by the Bayesian observers. This is mostly due to the generalization abilities of the function approximation. A learner using a table with entries for every state-action-reward mapping [48] has to effectively relearn its policy with every change in conditions.

Shift in Uncertainty Structure. We can also adapt our settings to simulate the conditions used in the experiment by Wallace and Stein [27] to introduce mismatches in the spatial alignment of stimuli from a common object. We ask whether reinforcement mediated learning could also produce results similar to the aberrant spatial integration found in their study. Therefore we bias the auditory signal by setting the mean of its noise distribution to a value different from zero.

Figure 9A shows contour plots of the Q-value function for one particular action after normal (filled) and biased training (empty). The area which favors integration (red) shifts by as many positions on the auditory axis as are introduced by the bias. The same is true for the unisensory tuning curves (Figure 9B), which were generated by plotting the response of the same output neuron to sequential single stimulation of each unisensory input neuron. These results are qualitatively similar to the ones reported by Wallace and Stein for the relationship between auditory and visual receptive fields of single neurons in cat superior colliculus.

Discussion

The fact that cue integration in sensory inference can be well matched by Bayesian models has led to the suggestion that such computations are implemented in the brain by explicit computations with uncertainties. Accordingly, current research is looking for ways in which populations of neurons could implement Bayesian computations involving probability distributions [16,22]. This view has led to the often implicit and sometimes explicit assumption, see e.g. [5,49,50], that reward dependent model free learning does not mediate this behavior. Often these investigations are accompanied by the implicit assumption, that a single algorithm has to be attributed to the brain, despite the fact that recent work by Daw and colleagues [51,52] has shown that learning in certain tasks can be best explained by assuming that multiple learning systems implementing different algorithms are working together in the brain (see also [53–56]). Currently it is unclear how cue integration and causal inference are learned over developmental timescales, as experiments with both children and animals suggest that cue integration abilities develop over time [17–20,23,27,57,58].

Bayesian models of cue integration have been extended to cases in which there is uncertainty about the underlying scene configuration that generated the sensory stimuli [12,13]. In this case of causal inference the observer has to judge how likely one of the possible scene configurations may have caused the observed sensory signals. There are two main ways how this could be done. Either the observer decides on the more likely scene layout and then interprets the signals according to this layout (model selection), or the positions of the objects in the scene are judged according to the likely positions according to both models and then are weighted by how much evidence there is for either layout (model averaging). Current empirical evidence is inconclusive whether human performance in such tasks is better explained by model selection, model averaging, probability matching or all of those [13,15,59,60]. Similarly, it is unclear, whether any additional contextual or task effects might affect the strategy used.

Here we investigated whether a reward based model free learner using function approximation is able to learn an orienting task requiring integration of cues. Furthermore, as the cues could

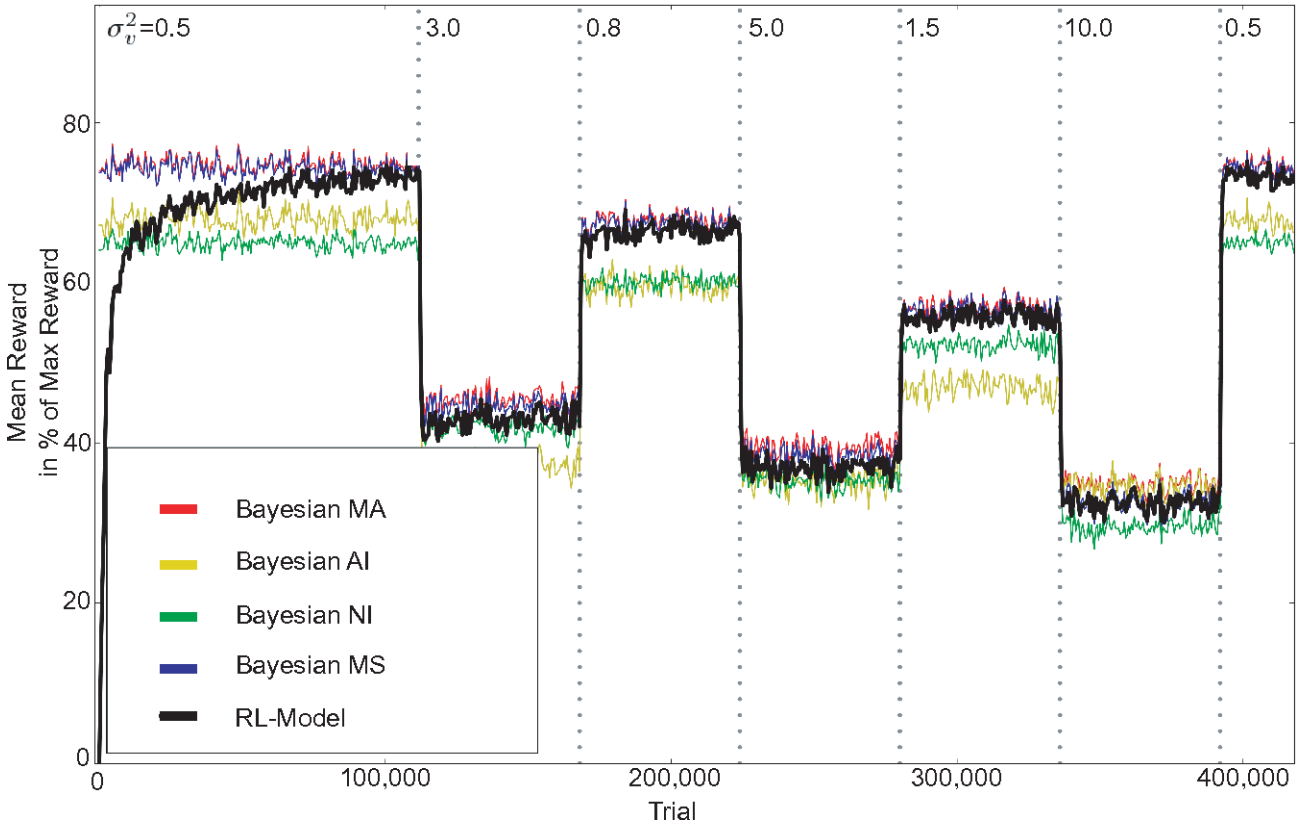

Figure 8. Performance of the RL Model for two outputs and temporally changing reliabilities. Reward obtained by the learner when choosing the action with highest predicted reward (black) compared to the different Bayesian observers. At each dotted vertical line visual reliability changes. Each data point is the sliding average of 1000 trials. $\sigma_a^2 = 3$.
doi:10.1371/journal.pone.0021575.g008

either originate from different sources or a single one, it was necessary to learn not to combine the estimates always, but to take into account that at larger separations of the two cues it is more likely that they originate from two different sources. The learner was given two audio-visual orienting tasks to solve. In the first task the learner was rewarded for orienting towards either one of the two stimuli, whereas in the second task the learner was rewarded separately for judging both the position of the visual and the auditory sources.

Under both task conditions the learner was able to carry out actions that combined cues according to their relative reliabilities. The reward obtained when following the reinforcement learner is higher than that obtained by the Bayesian learners that always or never integrate. It was also shown that the behavior of the RL model best matches that of a MA observer. This does not necessarily mean that humans always use MA, but shows the general ability of RL to approach optimal behavior. A recent paper by Wozny and colleagues [15] found evidence for a majority of subjects acting most similar to probability matching (at the causal inference level), but also a significant number of people that were better fit by model averaging. Further research is needed to clarify whether this is generally true or depends on additional parameters.

We could also show that the RL approach is able to deal with more complex uncertainty structures in the input. Here, the uncertainties are implicitly represented in the function approximation scheme of the value functions. Arguably, representing only uncertainties that are relevant for obtaining rewards is more economical than representing all potential distributions over all available scene variables. Indeed, here the distributions over sensory cues given relevant scene variables were not provided beforehand to the system, as is common in the Bayesian cue integration and causal inference setting. The proposed model was able to also perform similarly to the Bayesian predictions when the data likelihood was variable in time or space, when using non-uniform priors, and for changes in the causal structure. Humans were shown to be able to rapidly adapt to changes in cue reliability [61–63] and causal layout [64]. Although we do not want to claim that this is necessarily mediated by reward for the very early adaptation, we show the potential of RL-mechanisms to react to those changes. It would also be interesting to test children for the developmental aspects of such rapid re-weighting [65], but more experiments will be needed to clarify those results.

One feature that is missing in our approach is temporal relations between signals, which in a natural environment provide an important cue for causal inference (e.g. [66]). It was shown that this influence is also plastic in children [20] and in adults [67], so it would be interesting to see how reward mediated learning deals with the incorporation of temporal information. The TD–learning framework is in principle able to deal with delayed rewards. This question will be addressed by future work.

All learning was done with immediate reward feedback to individual actions using learning rules that have been well established in conjunction with reward related learning and orienting movements [31,32]. We are aware that using gradient descent learning to update the weights of the neural network could be considered problematic for a neural implementation [68]. In

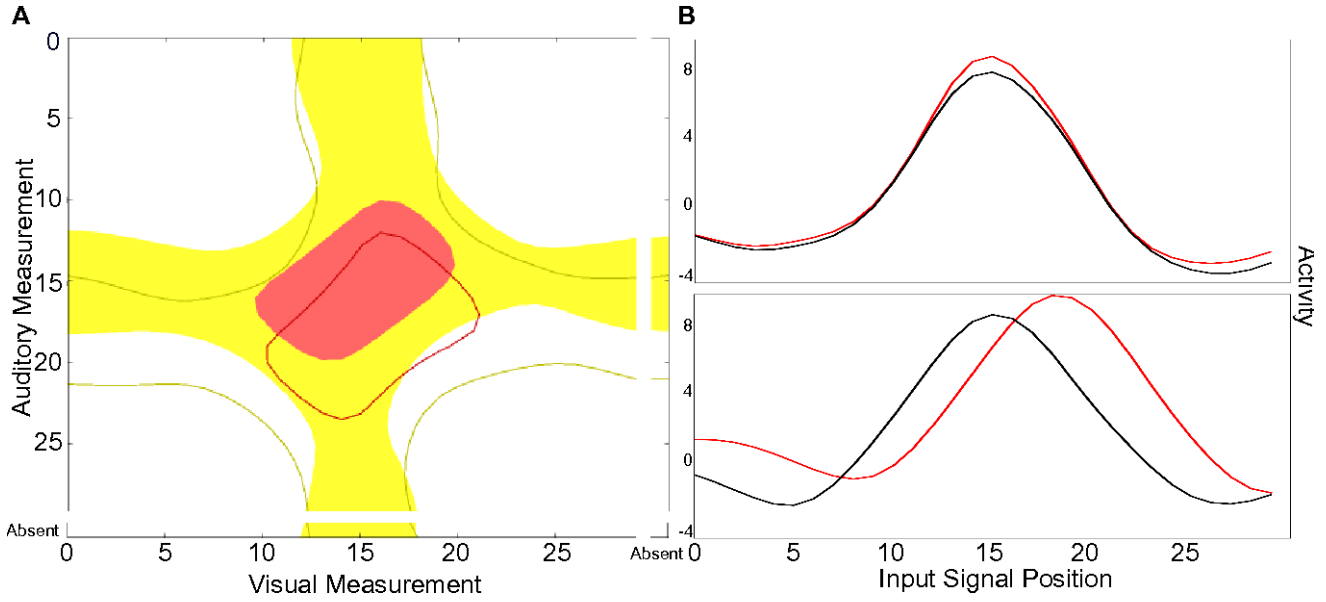

Figure 9. Responses of output neurons after training with auditory shift data. A: Overlay of contour plots of the Q-value functions for one action after unbiased training (filled areas) and after training with a 3-position shift in the mean of the auditory noise (empty areas). The contours include areas with predicted reward values higher than 10 (red) and 6 (yellow). B: Unisensory tuning curves of the same output neuron for biased (red) and unbiased conditions. The maximum visual response (top) does not change, whereas it is shifted by 3 positions in the auditory domain (bottom). ($\sigma_a^2 = \sigma_v^2 = 3$).
doi:10.1371/journal.pone.0021575.g009

the recent past, attempts were made to relate this kind of learning more closely to biology [69–71]. Future work will nevertheless try to use alternative solutions for learning of the synaptic weights.

Unfortunately we were not able to identify meaningful intermediate behavioral strategies while the model was still learning. It would be interesting to compare the behavior of the RL agent with recent empirical and theoretical work on the learning of cue integration, which suggest potentially different behavior such as calibration of a less reliable modality by a more reliable one [17,28,72,73] or using the modalities alternatingly [18] maybe according to the so called race model [19,20]. The modality providing the basis for calibration could depend on the relative reliabilities, could be innately determined or chosen randomly. Consistent with the first option are results showing that even unisensory performance in certain non–visual tasks can be worse in early blind compared to sighted children [72].

To conclude, the RL algorithm with function approximation was capable of learning near optimal performance in the Bayesian sense for both cue integration and causal inference tasks (consistent with previous results with tabular RL [48]). Importantly, despite not performing explicit computations with uncertainties, the reinforcement learner successfully changed actions depending on the uncertainty in the stimulus. Considerable evidence about the neural basis of such algorithms makes this approach appealing. Furthermore, it gives a direct way of accommodating learning of cue integration and causal inference over developmental timescales. Thus, even if RL algorithms may not be the only mechanisms underlying the human development of cue integration and causal inference they could definitely contribute to their development.

Methods

Task Setup

In our task each trial consists of the presentation of two stimuli in the visual and auditory modalities. These stimuli either originate from a single common source (Fig. 1 left) for the auditory and visual cue or from two separate sources/objects (Fig. 1 right). A position x along the spatial dimension is chosen from a uniform distribution for each object in the scene (but results if, e.g., visual prior is Gaussian around the central region are not different – Table 1 1C). In the two objects case we call their positions x_a (for the object that emits only an auditory signal) and x_v (for the object that emits only a visual signal). Space is discretized to $x_{\max} = 30$ positions for ease of computation. The received sensory signals are noisy versions of the true source locations. We use additive noise with normal distributions with zero mean and variances σ_a^2 and σ_v^2. Note that the RL model is also able to deal with noise from different distributions since we do not implement the learner based on a fixed distribution. See Table 1 2B for a setup with auditory and visual noise drawn from a logistic distribution with median 0. This noise is thought as being of sensory and/or environmental origin, e.g. background noise, neuronal firing stochasticities and tuning densities. Usually the variance of the auditory estimate is set larger than the visual one, in accordance with psychophysical observations for spatial tasks [74]. We call this noisy signal position z^a and z^v respectively. If the noise makes a signal fall outside of the spatial range, the stimulus is treated as not present, thus resulting in a unisensory training trial. An important implication of this setting is that the structure of observations is the same for both possible underlying generative models.

We use two slightly different versions of this task. In the single output task the learner has to orient towards a single location. That means in the case of two objects the reward only depends on the distance to the object closest to the estimated position. In the two outputs task it is required to orient towards both the visual and the auditory positions of their respective cause. In case of a common cause this should result in both estimates being equal. There are separate rewards for the visual and auditory action. The inputs were the same for both experiments.

Reinforcement Learning Model

An approximation of the function relating state-action pairs to predicted reward is learned. A three-layered neural network (see Fig. S1) is set up with an input unit for each position in every modality (here 60 input neurons). It should be mentioned that the yet unsolved problem of limited scalability of RL approaches for very large numbers of inputs, does also apply to our model. The input neurons i are all-to-all connected with weights $v_{i,j}$ to neurons j in the hidden layer (here j = 0…29). Stimulus locations z^a and z^v are represented by the population activity of these input neurons (see e.g. [75,76] for biological examples of population codes) in each modality separately (the first half of the neuron coding for the auditory input, the second for the visual one).

Each input neuron has a Gaussian receptive field, centered on position z_i^a or respectively $z_{i-x_{\max}}^v$ for $i \geq x_{\max}$. The variance of these Gaussians is in the order of the noise of the input stimuli. Overlapping receptive fields of the input neurons simply help the network to discover a spatial relationship between the possible input positions. We also tested the framework with simple binary input units and found no difference in the final results besides an increase in learning time (see Table 1 1B).

A sigmoidal transfer function on the sum of the weighted inputs g_i produces the activation y_j of the hidden neurons. These are again fully connected to the output neurons x' with weights $w_{j,x'}$. For every action x' there is one output unit, with its activation $o_{x'}$, given by the weighted sum of the hidden layer activity, representing an approximation of the appropriate Q-value. All weights are drawn from a uniform distribution, the v's between -0.1 and 0.1, the w's between -1 and 1.

Based on the network's outputs the learner chooses one of the available actions. This is done with the softmax function:

$$P(\hat{x}=x'|s)=\frac{e^{Q_{x',s}/\tau}}{\sum_{\bar{a}} e^{Q_{\bar{a},s}/\tau}}. \quad (1)$$

This probabilistic action selection rule chooses an action x' with a probability proportional to the relative predicted reward $Q_{x',s}$ for that action, given state s. We start with a high temperature parameter $\tau=\tau_0$, so that the learner chooses his actions only weakly influenced by the initial reward expectations. τ then decreases exponentially with learning time (with $\tau(t)=\tau_0^{\frac{v_\tau - t}{v_\tau}}$), passing 1 after a given number of steps v_τ. At smaller values of τ the selection favors more and more the action with highest expected reward, thus exploiting the environment.

After performing the selected action $\hat{x}$, the learner receives the true reward $r(\hat{x})$. We use a reward function that is maximal if $\hat{x}$ equals the true object position x, decaying quadratically with increasing distance within a surrounding area (with radius ρ) and zero otherwise.

For single output:

$$r(\hat{x}|x_a,x_v)=\max(0,(\rho-\min(|x_a-\hat{x}|,|x_v-\hat{x}|)))^2 \quad (2)$$

For two outputs:

$$\begin{aligned} r_a(\hat{x}_a|x_a,x_v)&=\max(0,(\rho-|x_a-\hat{x}_a|))^2,\\ r_v(\hat{x}_v|x_a,x_v)&=\max(0,(\rho-|x_v-\hat{x}_v|))^2 \end{aligned} \quad (3)$$

If only one object is present, the two position are equal, i.e. $x_a=x_v$. In the experiments shown above we used $\rho=4$. Changes of of ρ other than setting it to zero (only rewarding correct actions) only have an impact on the learning time. We also tested the model with an asymmetric reward function, where a correct visual action would only provide half the reward of a correct auditory action (results see Table 1 2C).

Based on the true reward, the Q-values for the particular state-action pair will be updated proportional to the difference between prediction $Q_{\hat{x},s}$ and $r(\hat{x})$. This difference can be seen as a temporal difference (TD) error for a single timestep. TD learning in general uses discounted future rewards for computing the prediction error: The Q-value function will not only represent the expected reward of a single state–action pair, but also include possible future rewards that are expected from the new state. In the present work the learner has to only perform a single action per trial and receives only immediate reward.

To minimize the TD error we use gradient descent to change the weights of the neural network by $\Delta w/\Delta v$ [77] with:

$$\Delta w_{j,x'}=\begin{cases} -\varepsilon(r_{\hat{x}}-o_{\hat{x}})(-y_j), & if \quad x'=\hat{x} \\ 0, & else \end{cases} \quad (4)$$

$$\Delta v_{i,j}=-\varepsilon(r_{\hat{x}}-o_{\hat{x}})(-w_{j,\hat{x}})y_j(1-y_j)g_i. \quad (5)$$

ε is an exponentially decreasing learning rate: $\varepsilon(t)=10^{\log(\varepsilon_0)-\frac{t}{v_\varepsilon}}$, with $\varepsilon_0=0.05$ and $v_\varepsilon=100{,}000$. The results did not change when using an alternative function for the learning rate, $\varepsilon(t)=\frac{\varepsilon_0}{\mathrm{ceil}(\frac{t}{v_\varepsilon})}$, with $v_\varepsilon=10{,}000$.

Bayesian Observer Models

We compare the performance of our model with that of four different Bayesian observers, inferring the position of the object given the input and the generating model C (Fig. 1 bottom). With Bayes' theorem and the assumption that the noise of different modalities is independent we can write the posterior probability as:

$$p(x|z^a,z^v)=\frac{p(z^a,z^v|x)p(x)}{p(z^a,z^v)}=\frac{p(z^a|x)p(z^v|x)p(x)}{p(z^a,z^v)}. \quad (6)$$

where the last equality is only valid if the two cues are conditionally independent given their cause. The likelihoods $p(z^a|x)$ and $p(z^v|x)$ include all information available from the input. The reliability of a cue is inversely proportional to the standard deviation of this distribution. In the experiments reported in this paper the prior $p(x)$ is always uniformly distributed. Other priors were used in simulations not shown, and the RL algorithm was able to adjust to these and still perform close to the Bayesian predictions. Since we are interested in the performance of the model in terms of reward, actions are not chosen only based on the posterior probabilities, but on the utility function $U(\hat{x}|z^a,z^v)$, which additionally takes into account the expected reward $r(\hat{x}_{[a,v]}|x_a,x_v)$ (we write $[a,v]$ to cover both the one and two output case) for a given action (see below). The use of different utility functions can accommodate different tasks in a very direct way and makes the behavioral goal explicit.

The Bayesian observers used here differ in the way they handle the two different possible generative models (one vs. two causes; Fig. 1 bottom). Model Averaging (MA) uses a utility function that is a weighted average of the inference results of each model. The weights are determined by the probability for one versus two

objects $p(C|z^a,z^v)$. This probability can again be computed from known distributions using Bayes formula similar to (6).

$$U(\hat{x}_{[a,v]}|z^a,z^v) = \begin{array}{c} p(C=1|z^a,z^v)\int\ r(\hat{x}_{[a,v]}|x_a,x_v)p(x|z^a,z^v)dx \\ +p(C=2|z^a,z^v)\int\int\ r(\hat{x}_{[a,v]}|x_a,x_v)p(x_a|z^a)p(x_v|z^v)dx_a dx_v \end{array} \quad (7)$$

.Model selection (MS) in contrast uses only the utility function of the most probable model.

$$U(\hat{x}_{[a,v]}|z^a,z^v) = \begin{cases} \int\ r(\hat{x}_{[a,v]}|x_a,x_v)p(x|z^a,z^v)dx, & if \quad p(C=1|z^a,z^v)>0.5 \\ \int\int\ r(\hat{x}_{[a,v]},x_a,x_v)p(x_a|z^a)p(x_v|z^v)dx_a dx_v, & else \end{cases} \quad (8)$$

We use a uniform prior over the number of objects in the scene ($P(C=1)=P(C=2)=0.5$). Results of additional simulations not shown here lead to similar results for asymmetric distributions.

We also consider two observers that only do inference on one model, ignoring the second one – one always integrates the information (AI) and the other always treats the inputs as independent (NI). The utility functions of all observer models are computed by numerical integration. For a given input we choose the action with maximum utility.

Another possible observer model would compute the same probability distributions as in MS and MA, but then select stochastically from them instead of choosing the maximum. Such a behavior is often called Probability Matching (PM). In our case it could be used in two ways. A recent paper proposed PM at the level of causal inference [15], an action will be chosen according to one of the generating models with the probability for that model to be the underlying cause ($P(C)$). Because this is an intermediate between MA and MS we only consider it when computing the R^2, where we distinguish between those. The second possibility would be to use PM for the action selection step, which was found in various studies to be a strategy employed by human observers in certain tasks [78,79]. This is actually implicitly assumed in our model by using the softmax function to pick the action, thereby we do not include this option in our analysis.

Supporting Information

Figure S1 Sketch of the neural network used for approximation of the Q-value function. Setup for the two–step orienting task, the setup for the simple orienting task differs only in that the network has only half as many output neurons, since only a single action is required.
(EPS)

Figure S2 PSEs of the model for the 2-AFC task. Plot of the PSEs for each of 5 repetitions of training and 2-AFC testing with different values for the variance of the visual noise σ_v^2. For all trials $\sigma_a^2=5$. In the test trials the visual signal of the standard was always at 18, the auditory at 12. PSEs of the RL model (black crosses) and Bayesian integration (red circles).
(EPS)

Figure S3 R^2 values of different observers for the responses of the RL model for all inputs. Same Plot as Fig. 6 but using inputs from the full space. The black and gray bars show the results for the auditory and the visual output for 50,000 trials. Mean over 10 training sessions with $\sigma_a^2=3$, $\sigma_v^2=2$, errorbars show standard deviation.
(EPS)

Figure S4 Input weights of the NN for the foveation setup. Input weights $v_{i,j}$ to representative hidden neurons. The left plot shows the weights only from visual input neurons ($i=[0:x_{\max}-1]$), the right only from the auditory input neurons ($i=[x_{\max}:2x_{\max}-1]$).
(EPS)

Author Contributions

Conceived and designed the experiments: THW CAR TR JT. Performed the experiments: THW. Analyzed the data: THW CAR JT. Wrote the paper: THW CAR TR JT.

References

1. Ernst MO (2004) Merging the senses into a robust percept. Trends in Cognitive Sciences 8: 162–169.
2. Kersten D, Yuille AL (2003) Bayesian models of object perception. Current Opinion in Neurobiology 13: 150–158.
3. Kersten D, Mamassian P, Yuille AL (2004) Object Perception as Bayesian Inference. Annual Review of Psychology 55: 271–304.
4. Knill DC, Richards W (1996) Perception as Bayesian inference. New YorkNew York, USA: Cambridge University Press.
5. Knill DC, Pouget A (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. Trends in Neurosciences 27: 712–719.
6. Rothkopf CA, Weisswange TH, Triesch J (2010) Computational modeling of multisensory object perception. New York: Springer, chapter 3. pp 21–53. doi:10.1007/978-1-4419-5615-6 3. URL http://www.springerlink.com/content/u45134054ktvl422/?p=3065735194a24778bf63cf25e7bcb01b&pi=2.
7. Alais D, Burr D (2004) The Ventriloquist E_ect Results from Near-Optimal Bimodal Integration. Current Biology 14: 257–262.
8. Battaglia PW, Jacobs RA, Aslin RN (2003) Bayesian integration of visual and auditory signals for spatial localization. Journal of the Optical Society of America A, Optics, image science, and vision 20: 1391–1397.
9. Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415: 429–433.
10. Knill DC, Saunders JA (2003) Do humans optimally integrate stereo and texture information for judgments of surface slant? Vision Research 43: 2539–2558.
11. Jacobs RA (1999) Optimal integration of texture and motion cues to depth. Vision Research 39: 3621–3629.
12. Sato Y, Toyoizumi T, Aihara K (2007) Bayesian inference explains perception of unity and ventriloquism aftere_ect: identi_cation of common sources of audiovisual stimuli. Neural Computation 19: 3335–3355.
13. Körding KP, Beierholm UR, Ma WJ, Quartz S, Tenenbaum JB, et al. (2007) Causal inference in multisensory perception. PLoS ONE 2: e943.
14. Shams L, Beierholm UR (2010) Causal inference in perception. Trends in Cognitive Sciences 14: 425–432.
15. Wozny DR, Beierholm UR, Shams L (2010) Probability Matching as a Computational Strategy Used in Perception. PLoS Computational Biology 6.
16. Fiser J, Berkes P, Orbán G, Lengyel M (2010) Statistically optimal perception and learning: from behavior to neural representations. Trends in Cognitive Sciences 14: 119–130.
17. Gori M, Del Viva M, Sandini G, Burr DC (2008) Young Children Do Not Integrate Visual and Haptic Form Information. Current Biology 18: 694–698.
18. Nardini M, Jones P, Bedford R, Braddick O (2008) Development of cue integration in human navigation. Current Biology 18: 689–693.
19. Nardini M, Bedford R, Mareschal D (2010) Fusion of visual cues is not mandatory in children. Proceedings of the National Academy of Sciences of the United States of America 107: 17041–17046.
20. Neil PA, Chee-Ruiter C, Scheier C, Lewkowicz DJ, Shimojo S (2006) Development of multisensory spatial integration and perception in humans. Developmental Science 9: 454–464.
21. Beck JM, Ma WJJ, Kiani R, Hanks T, Churchland AK, et al. (2008) Probabilistic population codes for Bayesian decision making. Neuron 60: 1142–1152.
22. Ma WJ, Beck JM, Latham PE, Pouget A (2006) Bayesian inference with probabilistic population codes. Nature Neuroscience 9: 1432–1438.
23. Putzar L, Goerendt I, Lange K, Rösler F, Röder B (2007) Early visual deprivation impairs multisensory interactions in humans. Nature Neuroscience 10: 1243–1245.
24. Schorr EA, Fox NA, van Wassenhove V, Knudsen EI (2005) Auditory-visual fusion in speech perception in children with cochlear implants. Proceedings of the National Academy of Sciences of the United States of America 102: 18748–50.
25. Wallace MT, Stein BE (1997) Development of multisensory neurons and multisensory integration in cat superior colliculus. Journal of Neuroscience 17: 2429–2444.
26. Wallace MT, Stein BE (2001) Sensory and multisensory responses in the newborn monkey superior colliculus. Journal of Neuroscience 21: 8886–8894.

27. Wallace MT, Stein BE (2007) Early experience determines how the senses will interact. Journal of Neurophysiology 97: 921–926.
28. Knudsen EI, Brainard MS (1991) Visual instruction of the neural map of auditory space in the developing optic tectum. Science 253: 85–87.
29. Hikosaka O, Nakamura K, Nakahara H (2006) Basal ganglia orient eyes to reward. Journal of neurophysiology 95: 567–84.
30. Takikawa Y, Kawagoe R, Itoh H, Nakahara H, Hikosaka O (2002) Modulation of saccadic eye movements by predicted reward outcome. Experimental Brain Research 142: 284–91.
31. Schultz W (2000) Multiple reward signals in the brain. Nature reviews Neuroscience 1: 199–207.
32. Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275: 1593–9.
33. Morris G, Nevet A, Arkadir D, Vaadia E, Bergman H (2006) Midbrain dopamine neurons encode decisions for future action. Nature Neuroscience 9: 1057–63.
34. D'Ardenne K, McClure SM, Nystrom LE, Cohen JD (2008) BOLD responses reecting dopaminergic signals in the human ventral tegmental area. Science 319: 1264–7.
35. Gottfried JA, O'Doherty J, Dolan RJ (2003) Encoding predictive reward value in human amygdale and orbitofrontal cortex. Science 301: 1104–7.
36. Haruno M, Kawato M (2006) Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. Journal of neurophysiology 95: 948–59.
37. O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, et al. (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 304: 452–4.
38. Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006) Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442: 1042–5.
39. Pessiglione M, Petrovic P, Daunizeau J, Palminteri S, Dolan RJ, et al. (2008) Subliminal instrumental conditioning demonstrated in the human brain. Neuron 59: 561–7.
40. Schönberg T, Daw ND, Joel D, O'Doherty JP (2007) Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. Journal of Neuroscience 27: 12860–7.
41. Sutton RS, Barto AG (1998) Reinforcement Learning: An Introduction. Cambridge, UK: MIT Press. 322 p. URL http://webdocs.cs.ualberta.ca/sutton/book/the-book.html.
42. Samejima K, Ueda Y, Doya K, Kimura M (2005) Representation of action-specific reward values in the striatum. Science 310: 1337–40.
43. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. Nature 441: 876–9.
44. Rangel A, Hare T (2010) Neural computations associated with goal-directed choice. Current opinion in neurobiology 20: 262–270.
45. Rao CR (1973) Linear Statistical Inference and its Applications. New York: Wiley-Interscience, 2nd edition. 656 p.
46. Hairston W, Wallace M, Vaughan J, Stein B, Norris J, et al. (2003) Visual localization ability inuences cross-modal bias. Journal of Cognitive Neuroscience 15: 20–29.
47. Knill DC (2005) Reaching for visual cues to depth: the brain combines depth cues differently for motor control and perception. Journal of Vision 5: 103–115.
48. Weisswange TH, Rothkopf CA, Rodemann T, Triesch J (2009) Can reinforcement learning explain the development of causal inference in multisensory integration? In:8th International Conference on Development and Learning, IEEE. pp 1–7. doi:10.1109/DEVLRN.2009.5175531. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber = 5175531.
49. Körding KP, Fukunaga I, Howard IS, Ingram JN, Wolpert DM (2004) A neuroeconomics approach to inferring utility functions in sensorimotor control. PLoS biology 2: e330.
50. Whiteley L, Sahani M (2008) Implicit knowledge of visual uncertainty guides decisions with asymmetric outcomes. Journal of Vision 8: 2.1–15.
51. Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nature Neuroscience 8: 1704–11.
52. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-Based Inuences on Humans' Choices and Striatal Prediction Errors. Neuron 69: 1204–1215.
53. Keramati M, Dezfouli A, Piray P (2011) Speed/Accuracy Trade-O_ between the Habitual and the Goal-Directed Processes. PLoS Computational Biology 7: e1002055.
54. Lengyel M, Dayan P (2008) Hippocampal contributions to control: the third way. In: Platt J, Koller D, Singer Y, Roweis S, eds. Advances in Neural Information Processing Systems 20. MIT Press. pp 889–896.
55. Shah A, Barto AG (2009) E_ect on movement selection of an evolving sensory representation: a multiple controller model of skill acquisition. Brain Research 1299: 55–73.
56. Sloman SA (1996) The Empirical Case for Two Systems of Reasoning. Psychological Bulletin 119: 3–22.
57. Brandwein AB, Foxe JJ, Russo NN, Altschuler TS, Gomes H, et al. (2011) The Development of Audiovisual Multisensory Integration Across Childhood and Early Adolescence: A High-Density Electrical Mapping Study. Cerebral Cortex 21: 1042–1055.
58. Held R, Ostrovsky Y, DeGelder B, Gandhi T, Ganesh S, et al. (2011) The newly sighted fail to match seen with felt. Nature Neuroscience 14: 551–553.
59. Helbig HB, Ernst MO (2007) Knowledge about a common source can promote visual- haptic integration. Perception 36: 1523–1533.
60. Stocker AA, Simoncelli EP (2008) A Bayesian Model of Conditioned Perception. In: Platt JC, Koller D, Singer Y, Roweis S, eds. Advances in Neural Information Processing Systems 20 (NIPS 2007). Vancouver, Canada: MIT Press. pp 1409–1416. URL http://books.nips.cc/papers/files/nips20/NIPS2007_1016.pdf.
61. Jacobs RA, Fine I (1999) Experience-dependent integration of texture and motion cues to depth. Vision Research 39: 4062–4075.
62. Triesch J, Ballard DH, Jacobs RA (2002) Fast temporal dynamics of visual cue integration. Perception 31: 421–434.
63. Young MJ, Landy MS, Maloney LT (1993) A perturbation analysis of depth perception from combinations of texture and motion cues. Vision Research 33: 2685–2696.
64. Wozny DR, Shams L (2011) Recalibration of Auditory Space following Milliseconds of Cross-Modal Discrepancy. Journal of Neuroscience 31: 4607–4612.
65. Bair WN, Kiemel T, Jeka JJ, Clark JE (2007) Development of multisensory reweighting for posture control in children. Experimental Brain Research 183: 435–446.
66. Lewald J, Guski R (2003) Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. Cognitive Brain Research 16: 468–478.
67. Fujisaki W, Shimojo S, Kashino M, Nishida S (2004) Recalibration of audiovisual simultaneity. Nature Neuroscience 7: 773–778.
68. Crick F (1989) The recent excitement about neural networks. Nature 337: 129–32.
69. D'Souza P, Liu SC, Hahnloser RHR (2010) Perceptron learning rule derived from spike-frequency adaptation and spike-time-dependent plasticity. Proceedings of the National Academy of Sciences of the United States of America 107: 4722–4727.
70. Roelfsema PR, van Ooyen A (2005) Attention-gated reinforcement learning of internal representations for classi_cation. Neural computation 17: 2176–214.
71. Tao HzW, Zhang LI, Bi Gq, Poo Mm (2000) Selective presynaptic propagation of long-term potentiation in de_ned neural networks. Journal of Neuroscience 20: 3233–43.
72. Gori M, Sandini G, Martinoli C, Burr D (2010) Poor Haptic Orientation Discrimination in Nonsighted Children May Reect Disruption of Cross-Sensory Calibration. Current Biology 20: 223–225.
73. Strelnikov K, Rosito M, Barone P (2011) Effect of Audiovisual Training on Monaural Spatial Hearing in Horizontal Plane. PLoS ONE 6: e18344.
74. Thomas GJ (1941) Experimental study of the inuence of vision on sound localization. Journal of Experimental Psychology 28: 163–177.
75. Fuzessery ZM, Wenstrup JJ, Pollak GD (1985) A representation of horizontal sound location in the inferior colliculus of the mustache bat (Pteronotus p. parnellii). Hearing research 20: 85–9.
76. Lee C, Rohrer WH, Sparks DL (1988) Population coding of saccadic eye movements by neurons in the superior colliculus. Nature 332: 357–60.
77. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning internal representations by error propagation. MIT Press. Volume 1 edition: 318–362.
78. Grant DA, Hake HW, Hornseth JP (1951) Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. Journal of experimental psychology 42: 1–5.
79. Rubinstein I (1959) Some factors in probability matching. Journal of experimental psychology 57: 413–6.