

---

# A Unifying Normative Framework of Decision Confidence

---

**Amelia M. Johnson**  
Department of Computer Science  
University of Washington  
Seattle, WA 98195  
ameliamj@uw.edu

**Micheal A. Buice**  
Allen Institute  
Seattle, WA 98109  
michaelbu@alleninstitute.org

**Koosha Khalvati**  
Allen Institute  
Seattle, WA 98109  
koosha.khalvati@alleninstitute.org

## Abstract

Self-assessment of one’s choices, i.e., confidence, is the topic of many decision neuroscience studies. Computational models of confidence, however, are limited to specific scenarios such as between choices with the same value. Here we present a normative framework for modeling decision confidence that is generalizable to various tasks and experimental setups. We further drive the implications of our model from both theoretical and experimental points of view. Specifically, we show that our model maps to the planning as an inference framework where the objective function is maximizing the gained reward and information entropy of the policy. Moreover, we validate our model on two different psychophysics experiments and show its superiority over other approaches in explaining subjects’ confidence reports <sup>1</sup>.

## 1 Introduction

Self-assessment of one’s choices, i.e., decision confidence, plays a key role in long-term decision-making and learning [1]. This assessment helps the decision maker improve their model of the outside world and consequently gain higher utility in the future [2, 3]. Due to this critical role, confidence has been the topic of many theoretical and experimental decision neuroscience studies. However, primarily focused on in perceptual decision-making experiments, confidence is mainly mathematically defined only for scenarios where different choices have the same potential reward. In these situations, based on some perceptual cues, the goal is to pick the “correct” choice instead of all other equally “incorrect” ones. Consequently, confidence is defined as the probability of choosing the correct option [1, 4]. In these setups, *perception confidence*, i.e., “what is the probability that my observation was correct?” and *decision confidence*, “what is the probability that I made the correct decision?” are inseparable because the choices do not differ in value.

In the real world, though, different actions lead to various potential rewards, and this variation in reward could influence confidence in decisions, differentiating it from perception confidence. A rustling sound is more likely to be because of the wind rather than a predator approaching. Still, everyone becomes more vigilant when hearing that sound as one of the possibilities is potentially life-threatening despite being unlikely. Moreover, everyone would be confident about the rationality

---

<sup>1</sup>The majority of this research was done while the first author was an intern at the Allen Institute.

of this decision. Notably, most experiments where different choices vary in value, known as value-based decision-making, do not involve uncertainty in perception [5]. These works are mostly about memory retrieval and valuation of different objects and do not study the interaction between value and perceptual uncertainty [6]. Here, we present a normative framework to formally define and assess decision confidence in a general scenario involving uncertainty about the outside world, prior knowledge of the decision maker about the world, and different utility functions for available choices. We model decision confidence as “probability of making the best decision”. In mathematical language, decision confidence is the probability of being optimal over a sequence of states and actions given the policy.

We further show that our approach to modeling subjects’ confidence in their decision equals to planning as inference framework [7, 8]. This framework maps to a reinforcement learning agent whose objective function is to jointly maximize the reward and the information entropy of the policy (also called maximum entropy reinforcement learning) [9, 10]. Moreover, we validate our framework by testing it on two different experiments on confidence evaluation and explaining its implications [11, 5].

## 2 Modelling Background and Problem Definition

Markov Decision Processes are graphical models used for optimal sequential decision-making in artificial agents. Recently, these frameworks have been applied successfully in modeling the behavior of subjects across various decision-making tasks and species in cognitive neuroscience [12, 13, 14, 15].

### 2.1 Fully and Partially Observable Markov Decision Processes

Formally, a Markov Decision Process (MDP) is a tuple  $(S, A, T, R)$  describing a Markovian system where  $S$  is the finite set of states,  $A$  is the finite set of actions,  $T = p(s'|a, s)$  is the transition function between states, and  $R$  is a bounded function representing the reward of each action in each state,  $r(s, a)$ . The goal of an MDP agent is to come up with the recipe of action selection, called *policy*, to maximize its total reward within *horizon*  $H$ . As the system is Markovian, each policy can be defined as a probability distribution of actions given the state, shown by  $\pi(a|s)$ . The optimal policy  $\pi^*$  is the policy that attains the goal of the agent:

$$\pi^* = \arg \max_{\pi} \mathbb{E}_{(s_t, a_t) \sim \pi(a_t|s_t)} \left[ \sum_{t=1}^H r(s, a_t) \right]. \quad (1)$$

The optimal policy can be obtained in polynomial time in the size of the state space, e.g., by using dynamic programming algorithms.

In most real-world situations, however, the environment is only partially observable, making the agent uncertain about the current state of the world. Partially Observable MDP (POMDP) models these situations by adding the concepts of observation, observation function, and the belief state to the MDP framework [16]. POMDP is formally defined as a tuple  $(S, A, Z, T, P, R)$  where  $S, A, T$  and  $R$  have the same definition as MDP.  $Z$  is the finite set of observations. In addition,  $P = p(z|s, a)$  is the observation function representing the probability of each observation, given the state and chosen action. A POMDP agent does not know the current state of the environment. Therefore, starting from a prior, called initial belief state  $b_1$ , it updates the posterior probability distribution over the states with each observation and action:

$$b_t(s) \propto p(z_t|s, a_{t-1}) \sum_{s' \in S} p(s|s', a_{t-1}) b_{t-1}(s'). \quad (2 \leq t \leq H) \quad (2)$$

In a POMDP, each policy can be represented as a mapping from belief states to a probability distribution over actions, i.e.,  $\pi(a|b)$  with the optimal policy  $\pi^*$  being the mapping that maximizes the expected total reward:

$$\pi^* = \arg \max_{\pi} \mathbb{E}_{(b_t, a_t) \sim \pi(a_t|b_t)} \left[ \sum_{t=1}^H \sum_{s \in S} b_t(s) r(s, a_t) \right] = \arg \max_{\pi} \mathbb{E}_{(b_t, a_t) \sim \pi(a_t|b_t)} \left[ \sum_{t=1}^H r(b_t, a_t) \right] \quad (3)$$

where  $r(b_t, a_t) = \sum_{s \in S} b_t(s) r(s_t, a_t)$  is the expected reward of belief state  $s_t$  if the agent choose action  $a_t$ .

POMDP can be viewed as a fully observable MDP with the state space of the belief state space of the original environment. Therefore, the optimal policy can be obtained in polynomial time in the size of the belief state. Since the belief state is a probability distribution over states, its space's size is exponential in the size of the state space. This means that transforming the POMDP to an MDP is not helpful unless the number of states is extremely low (less than 10) or the belief state can always be represented by a distribution of a few parameters. An example of the latter case is Kalman-filter-like environments where the belief state can always be represented with a Gaussian distribution with two parameters of  $\mu_t$  and  $\sigma_t$ .

## 2.2 Models of confidence in Cognitive Neuroscience

One of the applications of POMDPs and similar Bayesian frameworks is modeling the behavior “perceptual decision making” where the subject should select the “correct” choice based on some sensory observations to get reward [17]. As the term “correct” suggests, the reward function is symmetrical among different choices. In some of these studies, subjects also report their confidence in their choice. Numerous experiments have demonstrated that trained subjects perform similarly to an optimal agent such as POMDP. Moreover, their confidence in their choice closely matches the probability of choosing the correct option, i.e., the posterior probability of the most probable choice [1, 18] (or the sum of belief states that leads to the most probable choice). This close match is also called *Bayesian confidence hypothesis* [19]. Notably, due to reward symmetry in these situations, *perception confidence*, modeled as the belief about the hidden state, is inseparable from *decision confidence*, which would be the belief about the decision. More specifically, these experiments cannot flesh out the interactions between value, perception, and confidence and cannot test theoretical models for these interactions.

There are a few confidence experiments with asymmetries in the reward function. Models and methods of these studies, however, are all descriptive/statistical, e.g., positive correlation between confidence report and reward value [20, 21, 11]. Therefore, as opposed to *normative models*, these methods do not explain the reason behind the relationship of confidence with perception, prior, and rewards in a systematic and generalizable manner. Moreover, some experimental and theoretical works have studied “value-based confidence” in the context of “value-based decision making” where the task is to choose an object between offered options [22, 6, 5]. There is no perceptual ambiguity in these studies, and the variance in response is primarily due to the variability of the valuation of objects at different time points based on memory [6].

Finally, note that the optimality assumption in proposed models, such as the Bayesian confidence hypothesis, does not necessarily imply that the subject's confidence and performance are equal on average. As opposed to an AI POMDP agent, the subject is unaware of the exact generative functions of the environment ( $S$ ,  $T$ , and  $P$ ). The internal world model of the subject is built and learned by training. Therefore, the type and amount of training and feedback and the subject's capability and motivation in creating the accurate internal world model affect confidence. However, given the internal model, the behavior is optimal, e.g., the perception confidence follows the maximum belief state [18, 19]. We will explain this in more detail in our empirical evaluation and discussions.

## 2.3 Problem definition

We aim to formally define and test a normative framework that explains the interaction of perception and reward/value with the subject's confidence about their decision. Before presenting the model, we discuss a thought experiment as the motivation. This experiment consists of a set of trials, each of which begins with the presentation of a Gabor filter tilted left or right. The subject is asked to report their perception of the direction of the Gabor filter (left or right) and their confidence in their decision. Choosing the incorrect direction leads to no reward. Moreover, if correct, the direction left produces a reward 10 times the reward of the direction right. In an example trial, the subject believes the direction is right with a 0.7 probability. Assuming they want to gain a higher reward, they will choose the direction left due to the higher expected value, namely 0.3 versus 0.7. However, their perception confidence or “probability of choosing the correct/rewarding direction” is 0.3. On the other hand, their confidence in their decision, “likelihood of making the best decision” is probably

higher than 0.5 (because they chose it). Notably, while being higher than 0.5, their confidence in this trial is likely lower than in trials where they think the stimulus direction is left. However, from the strict optimality point of view, the decision confidence is 1 in both cases. Therefore, while a strictly optimal framework like POMDP offers a normative model of confidence (by explaining why confidence equals 1 through the lens of optimality), such a model is not aligned with reality.

Our goal is to systematically define and test decision confidence that includes all main aspects of decision-making, including reward, priors, and perceptual cues (observations) from the perspective of an optimal probabilistic decision-maker that is also aligned with reality. Notably, heuristics that combine reward and belief, such as the ratio of expected values of different actions, might work, but they are totally arbitrary and do not generalize to different tasks. We are looking for a *normative model* that explains *why* such a relationship exists between confidence and task parameters in a generalizable fashion.

### 3 Model

To formally model the decision confidence, we use the idea of *optimality* in the POMDP framework, where optimality is a binary variable that reflects receiving the maximum reward. The probability of optimality describes a subject's confidence or internal sense of whether their decision was optimal. This can be viewed as an agent choosing actions to maximize their total reward, modeled by a POMDP, and later evaluating the optimality of their decision.

Notably, the decision-making process might involve multiple actions. Therefore, what we refer to as “decision” is, in fact, a sequence of actions, given an observation after each action  $a_1, z_2, a_2, z_3, \dots, a_H$  which is called trajectory  $\tau$ . As the agent knows the observation function and the system is Markovian, this trajectory could also be expressed with belief states instead of observations:

$$\tau = a_1, b_2, a_2, b_3, \dots, a_H \quad (4)$$

Consequently, the probability of observing trajectory  $\tau$  in an optimal agent is  $p(\tau|O = 1, b_1)$ . This probability should be maximum for a trajectory that is generated by the optimal policy  $\pi^*$ .

With strict interpretation of optimality, only trajectories that are generated from the optimal policy  $\pi^*$  are optimal. This means that confidence is 1 for these trajectories and 0 for others. This “hard” definition of optimality punishes all non-optimal trajectories the same, meaning that a trajectory generated by a suboptimal policy is considered as non-optimal as a trajectory with the minimum possible outcome. This is not ideal—especially when the agents are humans, as humans are inherently suboptimal.

Given the fact that the system is Markovian and each policy is a mapping from belief states to actions, the probability of optimality can be expressed for each action given the belief state, i.e.,  $p(o_t = 1|a_t, b_t)$  (Fig. 1, left plot). With this representation, the probability of optimality for the whole trajectory can be expressed as  $p(o_{1:H} = 1|\tau)$ . Consequently, the probability of a trajectory being optimal is:

$$p(\tau|o_{1:H}) \propto p(\tau, o_{1:H}) = p(b_0) \prod_{t=0}^H p(o_t = 1|b_t, a_t) p(b_{t+1}|b_t, a_t). \quad (5)$$

If the belief state dynamics is deterministic, i.e., each action in each state leads to exactly one state and one observation, we would have:

$$p(\tau|o_{1:H}) \propto \mathbb{I}(p(\tau) \neq 0) \prod_{t=0}^H p(o_t = 1|b_t, a_t). \quad (6)$$

A high probability of optimality for a given belief state and action pair should reflect a higher reward for that pair of belief state and action (as higher optimality denotes a higher reward). Optimality should then be defined so that the total probability of optimality reflects the total reward sum. To transit between this total probability of optimality product and total reward sum, the probability of optimality for a given belief state and action is set to the exponentiation of the reward of the same belief state and action:

$$p(o_t|b_t, a_t) = e^{r(b_t, a_t)} \quad (7)$$

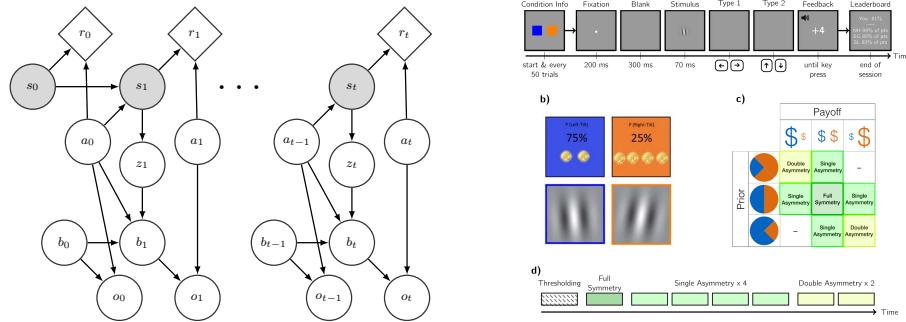


Figure 1: Left: Graphical model of the framework that measures the probability of making the optimal decision as confidence. Right: Experimental setup of a perceptual decision making task with varying prior and reward distribution (picture from [11]).

To make  $p(o_t|b_t, a_t)$  between 0 and 1, we can make all rewards negative by subtracting the maximum possible reward from them.

Notably, the above choice still looks arbitrary as there are many other functions that satisfy the constraints of our problem. Beyond being an intuitive way to relate a probability and a sum, if optimality is defined for an agent as it is in equation 7, then the agent will jointly maximize the total reward and the information entropy of the policy, which is also why our framework is a *normative model*. This result can be shown by deriving the policy from the optimal trajectory as defined above. One way to derive such a policy is to approximate  $p(\tau|o_{1:H})$ . If the approximation of  $p(o_t = 1|a_t, b_t)$  is expressed with policy  $\pi(a_t|b_t)$ , the approximation of optimal trajectory will be:

$$\hat{p}(\tau) = p(b_0) \prod_{t=0}^H p(b_{t+1}|b_t, a_t) \pi(a_t|b_t). \quad (8)$$

and the desired policy  $\pi(a_t|b_t)$  can be obtained by minimizing the KL-divergence (maximizing negative KL-divergence) between the approximate and true distributions:

$$\begin{aligned} -DL(\hat{p}(\tau)||p(\tau|o_{1:T})) &= \mathbb{E}_{\tau \sim \hat{p}(\tau)} \left[ \sum_{t=0}^H r(b_t, a_t) - \log \pi(a_t|b_t) \right] \\ &= \sum_{t=0}^H \mathbb{E}_{(b_t, a_t) \sim \hat{p}(b_t, a_t)} [r(b_t, a_t) + \mathcal{H}(\pi(a_t|b_t))] \end{aligned} \quad (9)$$

This derivation is known as “planning as inference” in the literature [7, 8], and equivalent of soft Q-learning on the belief state [9, 10]. Moreover, these equations also show that our definition of the probability of optimality works even when the dynamics of the belief state are stochastic.

### 3.1 Interpretation

According to our model, the agent makes decisions strictly optimally, like a POMDP. Its evaluation of optimality, however, allows for other trajectories through the concept of soft optimality. We only derived that policy of confidence judgment to demonstrate better why exponentiation of reward (equation 7) is a reasonable choice. This self-assessment is, to some degree, similar to inverse reinforcement learning. Notably, entropy regularization has been proven to be a practical approach in inverse reinforcement learning too [9]. In this approach, the evaluator allows some suboptimality to consider noises and hidden information. Such allowance in our confidence model considers imperfect learning of the environment or possible trial-to-trial changes in the computations. For example, arousal level affects perception and, consequently, performance. While in each trial, the most rewarding action should be chosen based on the received information, the evaluation should consider variation in the quality of this information gathering, e.g., “What if this decision was made under a low arousal level?” Maximum Entropy regularization is an agnostic approach to these variations and confounding factors.

Notably, the regularization usually accompanies a parameter  $\beta$  with a positive value, and  $p(o_t|b_t, a_t) = e^{\beta(b_t, a_t)}$  ( $\beta \geq 0$ ) still reflects the monotonic relationship between confidence (probability of being optimal) and the accumulated reward. We did not use that parameter in our equations for simplicity, and the fact that such an extra free parameter did not improve our fits to experimental data. However, with the extra parameter  $\beta$ , this model also include the hard optimality as  $\beta \rightarrow \inf$ .

### 3.2 Using the model on experimental data

Our framework proposes a confidence model for any sequential decision-making task under uncertainty. Specifically:

$$\text{confidence} \propto \prod_{t=1}^H p(o_t|b_t, a_t) = \prod_{t=1}^H e^{\mathbb{E}_{s_t}[r(s_t, a_t)]} \quad (10)$$

when the trajectory of the subject is  $\tau = a_1, b_2, a_2, b_3, \dots, a_H$  in a given trial. Fitting this equation to a subject's reported confidence could be extremely computationally expensive. One of the main reasons behind such computational cost is the belief state, with a space size exponential in the original state space size in the most general cases. Another reason is the need for normalization, which requires the calculation of the probability of optimality for all possible trajectories. In practice, however, fitting confidence is not intractable due to the simplicity of models of perception and the existing experimental setups. First, behavioral and even neural data on perception have been shown to be accurately modeled with Gaussian distributions. Second, in the current experimental setups, the dynamics of the task (transition function) are often very simple, and the number of actions is very minimal, e.g., one action of choosing one of the options after a perceptual cue. As a result, the belief state can always be expressed with a Gaussian distribution with two parameters of mean and variance. In other words,  $b_t$  could often be replaced with two one-dimensional parameters of  $\mu_t$  and  $\sigma_t^2$ . Finally, if the task has only one step of action selection, which is often the case, the confidence is simply proportional to  $e^{\mathbb{E}_s[r(s, a)]}$  where  $a$  is the chosen action. While this is convenient in terms of fitting, it brings another challenge. Any monotonic function of reward as the probability of optimality aligns it with the maximum total reward concept (instead of the exponential we used in equation 7 to map summation to product). For example, one convenient heuristic is the ratio of the expected reward of different actions, i.e., confidence  $\propto \mathbb{E}_s[r(s, a)]$ . Although this is arbitrary, we tested this intuitive definition of confidence, called "expected value ratio", and compared it to our definition of confidence on experimental data.

## 4 Results

We tested our model in two experiments, each focusing on different aspects and potential issues of confidence modeling. The first experiment was a perceptual decision-making task with asymmetric priors and rewards [11]. We focused on the interaction of reward and perception with confidence in our fits and analyses. The second experiment was a value-based decision-making task with no perceptual ambiguity [23], in which we tested our definition of optimality and compared our results with "expected value ratio" as a confidence hypothesis. To make our results more readable, we call our confidence model, "soft optimality" decision confidence.

### 4.1 Perceptual Decision Making Task with varying priors and rewards

**Experiment:** This experiment was designed to study the interaction of reward and priors with perceptual confidence [11]. In this experiment, 10 subjects were shown a Gabor filter tilted left or right and asked to report their perception of the direction of the Gabor filter (left or right) and, subsequently, their confidence in their perception (low or high) (as seen in Fig. 1, right plot). The difficulty of the trials (the extremity of the orientation of the Gabor filters) was constant across trials and was fitted to each subject before the main experiment to have approximately 70% accuracy. Each subject completed different sessions of this task where the prior probability distribution and the reward distribution for correction choices across the two directions were varied (prior probability was either 3:1, 1:1, or 1:3; payoffs were either 2:4, 3:3, or 4:2). The subjects were told the exact prior and reward distribution before each session and every 50 trials during the session. Each session was completed on separate days, with the fully symmetric trials being completed first, followed by the asymmetric trials in a random order. Each session consisted of 700 trials, the first 100 of which were

Table 1: AIC values for the fit to all confidence models in perceptual decision-making task

Model	1	2	3	4	5
Perception	1350.81 $\pm$ 42.17	<b>1668.53 <math>\pm</math> 2.98</b>	1763.63 $\pm$ 16.18	<b>1302.93 <math>\pm</math> 15.63</b>	<b>1441.89 <math>\pm</math> 3.19</b>
Soft Optimality	<b>1199.72 <math>\pm</math> 0.38</b>	1819.12 $\pm$ 20.12	<b>1674.94 <math>\pm</math> 3.02</b>	2056.63 $\pm$ 40.63	2309.59 $\pm$ 32.36
Observation	1316.56 $\pm$ 22.86	<b>1675.12 <math>\pm</math> 6.84</b>	<b>1673.77 <math>\pm</math> 3.03</b>	<b>1302.93 <math>\pm</math> 15.63</b>	<b>1441.89 <math>\pm</math> 3.19</b>
Expected Value Ratio	1346.6 $\pm$ 40.16	<b>1668.48 <math>\pm</math> 1.82</b>	1695.27 $\pm$ 7.89	<b>1295.91 <math>\pm</math> 13.51</b>	<b>1441.34 <math>\pm</math> 2.37</b>

Model	6	7	8	9	10
Perception	1383.24 $\pm$ 6.86	<b>1670.72 <math>\pm</math> 5.89</b>	1948.82 $\pm$ 52.65	1777.82 $\pm$ 32.36	1368.61 $\pm$ 1.29
Soft Optimality	1827.91 $\pm$ 22.73	1889.75 $\pm$ 20.14	<b>1655.49 <math>\pm</math> 10.5</b>	<b>1656.61 <math>\pm</math> 6.59</b>	<b>1358.02 <math>\pm</math> 1.07</b>
Observation	<b>1368.84 <math>\pm</math> 3.38</b>	<b>1682.17 <math>\pm</math> 12.09</b>	1748.28 $\pm$ 10.26	<b>1673.23 <math>\pm</math> 18.8</b>	1366.24 $\pm$ 1.47
Expected Value Ratio	<b>1376.47 <math>\pm</math> 4.98</b>	<b>1673.96 <math>\pm</math> 6.71</b>	1875.34 $\pm$ 42.0	1723.66 $\pm$ 25.62	1384.2 $\pm$ 1.99

Table 2: Rate/Probability of each reporting high confidence in each subject and each model's prediction for the trials with value asymmetry in the perceptual decision making task

Model	1	2	3	4	5	6	7	8	9	10
Experiment	80.28	53.16	49.33	21.96	28.37	25.37	46.84	57.65	42.43	75.87
Perception	60.78	48.58	37.17	16.22	26.93	17.57	51.13	29.77	26.34	68.1
Soft Optimality	79.29	68.72	52.74	61.96	69.84	55.78	68.42	50.17	47.0	69.5
Observation	63.35	55.52	65.06	16.22	26.93	23.34	55.99	43.62	38.82	68.43
Expected Value Ratio	61.06	49.17	43.4	16.89	27.42	18.82	51.76	32.77	29.69	83.87

discarded from analysis. Subjects received rewards of \$0-\$20 based on their performance. Notably, the subjects were explicitly instructed to report their confidence in the direction of the stimulus. In other words, they were explicitly requested to report their perception confidence. More details can be found in the original paper of this study [11].

**Fitting and Comparison:** First, we built each subject's model with a POMDP based on their choices. This POMDP contained 2 (hidden) states, each representing one of the directions. The perception of the subject (observation function) was modeled as a Gaussian distribution. Observations came from  $\mathcal{N}(-1, \sigma_z^2)$  and  $\mathcal{N}(1, \sigma_z^2)$  from direction left and right respectively. Notably, the actual generative process of observations could be different from the learned model by the subject [18]. Therefore, while observations were sampled from  $\mathcal{N}(\pm 1, \sigma_z^2)$ , the subject's internal model was  $\mathcal{N}(\pm 1, \sigma_{sz}^2)$ .  $\sigma_z^2$  is called external observational noise.  $\sigma_{sz}^2$  is the internal observational noise. Using the POMDP framework, the choice in each trial was obtained by sampling from the true generative process (with external noise), updating the belief based on the prior and learned observation function (with internal noise), and finally, picking the direction with highest expected reward by combining belief and the reward of each direction.

The external observational noise was first fit to each subject's choices in symmetric trials through gradient descent. Only in asymmetrical trials would the internal observational noise impact the subject's choice. Therefore, the internal observational noise was fit to the subject's choices in a subset of the trials with prior distribution asymmetries using a grid search and a maximum likelihood estimation with a Bernoulli likelihood function. The prior distribution was set to the actual prior value communicated to the subjects. Therefore, our model had only two free parameters in the fitting process, i.e., internal and external noise.

Based on the noise parameters obtained from the choice data, we could predict the subject's confidence in each trial according to different confidence models. The main two models were soft optimality confidence (ours) and perception confidence, which is the belief about the choice. We also included two more models. One was the observation likelihood (belief but without considering the prior), which we called observation confidence. The other model was the expected value ratio, as discussed before. For each confidence model, the confidence criterion, the threshold at which confidence is binarized into "low" or "high", was fit to the subjects' confidence reports in trials with asymmetric prior and fully symmetric trials (same subset of trials that we obtained noise parameters from). This fitting process included a grid search and a maximum likelihood estimation with a Bernoulli likelihood function.

Table 3: AIC values for the fit to all confidence models with an additional parameter accounting for the choice bias

Model	1	2	3	4	5
Perception	1364.92 $\pm$ 32.82	<b>1670.7 <math>\pm</math> 3.61</b>	1774.34 $\pm$ 21.38	<b>1301.69 <math>\pm</math> 12.47</b>	<b>1443.22 <math>\pm</math> 4.41</b>
Soft Optimality	<b>1201.72 <math>\pm</math> 0.65</b>	1818.99 $\pm$ 16.09	<b>1676.09 <math>\pm</math> 4.47</b>	2072.39 $\pm$ 42.81	2336.31 $\pm$ 77.72
Expected Value Ratio	1364.15 $\pm$ 33.0	1674.23 $\pm$ 2.88	<b>1669.71 <math>\pm</math> 12.56</b>	<b>1315.26 <math>\pm</math> 17.07</b>	<b>1433.92 <math>\pm</math> 11.62</b>

Model	6	7	8	9	10
Perception	<b>1386.16 <math>\pm</math> 12.54</b>	<b>1699.3 <math>\pm</math> 31.19</b>	1938.9 $\pm$ 41.8	1767.18 $\pm$ 19.07	1371.14 $\pm$ 2.06
Soft Optimality	1828.37 $\pm$ 19.27	1972.15 $\pm$ 93.6	<b>1657.1 <math>\pm</math> 7.11</b>	<b>1659.87 <math>\pm</math> 5.6</b>	<b>1360.43 <math>\pm</math> 1.84</b>
Expected Value Ratio	<b>1379.42 <math>\pm</math> 9.61</b>	<b>1669.33 <math>\pm</math> 23.49</b>	1866.07 $\pm$ 29.56	1709.11 $\pm$ 16.48	1385.87 $\pm$ 3.81

We tested the four mentioned models on the trials that had an asymmetry in the value distribution, which were not used to fit any of the parameters of these models. Notably, the model's values/rewards of actions were set to the values communicated to the subjects and were not free parameters. The AIC value ranges in table 1 are the results of fitting the model 10 times on randomly shuffled trials. The bolded values are the lowest AIC scores for a given subject, factoring in the range of uncertainty found in the average process across 10 trials. Moreover, table 2 shows the rate/probability of high-confidence reports by each subject and the prediction of the models in these trials (asymmetrical values; averaged over the 10 runs). Comparing these predictions to the actual high-confidence rate is not statistically reliable but more intuitive to humans.

As shown in table 1, our decision confidence model (soft optimality) was a better fit for five subjects compared to the perceptual confidence model. This is especially important because the subjects were explicitly instructed to report their perception confidence, yet half reported their confidence in choice. To be more precise, half could not override the assessment mechanism of their decisions. Importantly, adding the other two models (observation and expected value ratio) did not change our result. In fact, they strengthened our claim. The AIC value of our model often differs more significantly from the other three models. In other words, when our model performs better, it usually outperforms others significantly (and vice versa). This shows that the phenomenon we are observing is not an artifact of misuse of statistical tests (e.g., test's assumptions not holding) or fitting a powerful function.

We also got the same results while considering a choice bias in subjects towards one of the directions (as seen in table 3). We modeled this bias by adding a free parameter to the prior of each subject. We did not include the observation model as a comparison for these results because the observation model does not make use of a subject's prior distribution in its confidence computation. All subjects' behavior that our model fit best originally were also best explained by our model after the addition of the bias parameter.

## 4.2 Value Based Decision Making Task

To further test our soft-optimality assumption, especially compared to the expected value ratio heuristic, we applied our model to value-based decision making experiment [23].

**Experiment:** In this experiment, 33 subjects first did a rating task where they were shown pictures of different foods in succession and were asked to rate, on a sliding scale, how much they would want to eat that food at the end of the experiment (as seen in Fig. 2, left plot). After rating each food once, they were shown each food again and asked to rate it a second time to account for noise in their subjective assessment of the food's value. They were not told they would have to rate each food twice, so this didn't affect their initial rating of the food or allow them to try to memorize their first rating when rating a food for the second time. Subsequently, they were asked to do a decision-making task where they were presented with pictures of two of the foods they had previously rated and had to select which of the two they would rather eat and rate their confidence in their selection on a sliding scale. A "correct" decision was defined as a decision in which the subject chose the food object that they had rated higher on average between the two rating trials. Further task details are available in the original paper [23].

**Fitting and Comparison:** To generate predictions of subjects' confidence for the trials in this experiment, we used their average rating of a food during the rating trials to be an estimate of their



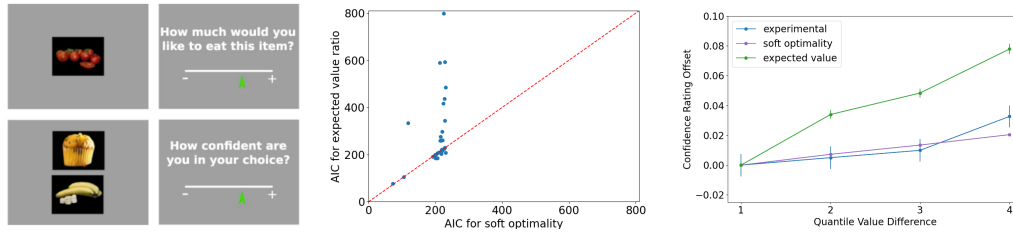


Figure 2: Left: Experimental setup of a value-based decision task with rating and decision trials (picture from [23]). Middle: Visual demonstration of AIC values for the soft optimality model versus the expected value ratio model on the value-based decision making task. The red dashed line represents equal AIC values. Right: As the quantile value difference increases, soft optimality predicts a rate of confidence increase closer to the experimental rate as compared to the rate predicted by the expected value method.

Table 4: AIC values by subject for different confidence prediction models on the value-based decision making task

Model	1	2	3	4	5	6	7	8	9	10	11
Expected Value Ratio	258.87	<b>224.61</b>	<b>219.06</b>	589.85	<b>210.31</b>	333.47	206.60	<b>202.14</b>	<b>190.56</b>	799.15	219.69
Soft Optimality	<b>216.79</b>	225.63	220.58	<b>212.84</b>	210.84	<b>118.28</b>	<b>206.40</b>	218.32	191.94	<b>224.71</b>	<b>219.48</b>

Model	12	13	14	15	16	17	18	19	20	21	22
Expected Value Ratio	<b>185.37</b>	261.71	297.92	220.79	223.33	<b>215.82</b>	260.02	417.31	<b>199.57</b>	229.26	199.56
Soft Optimality	200.20	<b>221.03</b>	<b>219.84</b>	<b>219.07</b>	<b>223.21</b>	218.58	<b>214.67</b>	<b>223.19</b>	200.84	<b>228.74</b>	<b>199.30</b>

Model	23	24	25	26	27	28	29	30	31	32	33
Expected Value Ratio	<b>207.01</b>	<b>215.36</b>	<b>104.43</b>	343.86	220.60	<b>185.43</b>	484.97	436.96	276.38	76.54	593.16
Soft Optimality	230.68	219.77	105.06	<b>228.51</b>	<b>220.16</b>	206.52	<b>230.81</b>	<b>227.34</b>	<b>215.55</b>	<b>71.99</b>	<b>228.66</b>

expected value of reward for a given option during the choice tasks. We then used this expected value of the belief state (where the belief state is their hypothesis as to which of the two choices is the “correct” choice) in our soft optimality equation and normalized this value to map our confidence prediction to the sliding scale that subjects use to report their confidence in their decision.

To assess our model’s ability to predict subjects’ confidence reporting behavior, we compared the fit of our soft optimality model to the fit of the expected value ratio model (as seen in Table 4 and Fig. 2, middle plot). Notably, since we used the values obtained from the rating section, none of the models have any free parameters. The additional two models used for comparison in the perceptual decision-making experiment couldn’t be used as metrics of comparison in this task as there was no perceptual component, which is a requirement of using the observation and perception models. On average, our soft optimality outperformed the expected value model with an average AIC value of 206.65 across all subjects as compared to an average AIC value of 282.11. The soft optimality model was also a better fit for the majority of the subjects; it best explained 21 out of the 33 subjects’ confidence reporting behavior. Moreover, the difference in AIC values is significantly larger in subjects whose behavior is better explained by our model (soft optimality) as demonstrated in Fig. 2, middle plot. We also compared the relationship between the difference in the value of the offered choices and the subject’s confidence (relative to their minimum confidence level, i.e. confidence offset, for normalization across all subjects) with the predictions of the two models. Our method better predicts the rate at which the subjects’ confidence will correspondingly increase (Fig. 2, right plot). Validating our model on this data set strengthens our conviction that it can generalize across multiple types of common decision paradigms and that it can be descriptive of more subjects than other comparable Bayesian approaches. Moreover, it further confirms the practicality of the exponential function in explaining confidence as opposed to heuristics like the expected value ratio.

## 5 Discussion

We present a normative framework that measures “decision confidence” and captures its interaction with reward, perception, and prior. Our model is essentially planning as inference framework, developed and discussed before [7, 10]. Our main contribution here is connecting this framework to the confidence judgment in humans. Notably, soft reinforcement learning has been used to fit subjects’ choices in decision-making before. However, those were actual decisions in the context of exploration-exploitation or imperfect rationality. We used this approach in action evaluation. In fact, according to our model, actions themselves follow strict optimality (modeled by a POMDP). We validated our results on two different experiments from different groups. Both datasets are publicly available and accessible in their corresponding papers [11, 23]. Our analysis code is also available at <https://github.com/ameliamj/decision-confidence-model>.

Current experiments, mostly including only 2 choices and 1 step of action selection, are insufficient to flesh out different aspects of confidence and its models. We believe the generalizability of our framework makes it a great candidate for testing and modeling confidence in more complicated setups. Future experiments could go in multiple directions. One example is confidence assessment in a sequence of actions instead of one action. One important aspect of our model, which is not present in other models and even most experiments, is the ability to assess confidence in a sequence of actions (trajectory). Multiple common experimental setups in the field, such as two-step task for humans [24] and even maze navigation for rodents can be used to study confidence of a sequence of actions (whole decision) if confidence assessment is added to them. Another example direction is confidence assessment in multiple choices. Two-choice tasks are too simple to distinguish between various models. Therefore, some researchers have started focusing on confidence assessment on multiple choices. For example, one study has shown that the difference between the probability of the top two choices explains confidence better than the posterior probability of the most likely state when three choices are presented [5]. In that study, the reward was the same for all options. It could also be illuminating to test confidence assessment when the rewards of different choices are not equal.

Confidence expression is an important factor in society and affects individuals’ interactions. Underprivileged individuals often suffer from low confidence, which may further hurt them in their role in their community. Therefore, we believe accurate estimation of confidence and comparing it with actual performance will reduce the societal gaps and promote fairness. However, computational models of confidence and, in general human behavior are very recent. Therefore, no impactful decision should be made based on these models (e.g., in job applications) until they are tested extensively in numerous controlled setups.

## References

- [1] Alexandre Pouget, Jan Drugowitsch, and Adam Kepecs. Confidence and certainty: distinct probabilistic quantities for different goals. *Nature Neuroscience*, 19(3):366–374, March 2016. Number: 3 Publisher: Nature Publishing Group.
- [2] Roozbeh Kiani and Michael N. Shadlen. Representation of Confidence Associated with a Decision by Neurons in the Parietal Cortex. *Science*, 324(5928):759–764, May 2009. Publisher: American Association for the Advancement of Science.
- [3] Jan Drugowitsch, André G Mendonça, Zachary F Mainen, and Alexandre Pouget. Learning optimal decisions with confidence. *Proceedings of the National Academy of Sciences*, 116(49):24872–24880, 2019. Publisher: National Acad Sciences.
- [4] Koosha Khalvati, Saghar Mirbagheri, Seongmin A. Park, Jean-Claude Dreher, and Rajesh PN Rao. A Bayesian Theory of Conformity in Collective Decision Making. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.
- [5] Hsin-Hung Li and Wei Ji Ma. Confidence reports in decision-making with multiple alternatives violate the Bayesian confidence hypothesis. *Nature Communications*, 11(1):2004, April 2020. Number: 1 Publisher: Nature Publishing Group.
- [6] Michael N. Shadlen and Daphna Shohamy. Decision Making and Sequential Sampling from Memory. *Neuron*, 90(5):927–939, June 2016.

- [7] Marc Toussaint and Amos Storkey. Probabilistic inference for solving discrete and continuous state Markov Decision Processes. In *Proceedings of the 23rd international conference on Machine learning - ICML '06*, pages 945–952, Pittsburgh, Pennsylvania, 2006. ACM Press.
- [8] Matthew Botvinick and Marc Toussaint. Planning as inference. *Trends in Cognitive Sciences*, 16(10):485–488, October 2012.
- [9] Brian D. Ziebart, Andrew Maas, J. Andrew Bagnell, and Anind K. Dey. Maximum entropy inverse reinforcement learning. In *Proceedings of the 23rd national conference on Artificial intelligence - Volume 3, AAAI'08*, pages 1433–1438, Chicago, Illinois, July 2008. AAAI Press.
- [10] Sergey Levine. Reinforcement Learning and Control as Probabilistic Inference: Tutorial and Review, May 2018. arXiv:1805.00909 [cs, stat].
- [11] Shannon M. Locke, Elon Gaffin-Cahn, Nadia Hosseinizadeh, Pascal Mamassian, and Michael S. Landy. Priors and payoffs in confidence judgments. *Attention, Perception, & Psychophysics*, 82(6):3158–3175, August 2020.
- [12] P. Dayan and N. D. Daw. Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience*, 8(4):429–453, December 2008.
- [13] Rajesh P. N. Rao. Decision Making Under Uncertainty: A Neural Model Based on Partially Observable Markov Decision Processes. *Frontiers in Computational Neuroscience*, 4:146, November 2010.
- [14] Koosha Khalvati and Rajesh PN Rao. A Bayesian Framework for Modeling Confidence in Perceptual Decision Making. In *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc., 2015.
- [15] Armin Lak, Emily Hueske, Junya Hirokawa, Paul Masset, Torben Ott, Anne E Urai, Tobias H Donner, Matteo Carandini, Susumu Tonegawa, Naoshige Uchida, and Adam Kepecs. Reinforcement biases subsequent perceptual decisions when confidence is low, a widespread behavioral phenomenon. *eLife*, 9:e49834, April 2020. Publisher: eLife Sciences Publications, Ltd.
- [16] Anthony R. Cassandra, Leslie Pack Kaelbling, and Michael L. Littman. Acting optimally in partially observable stochastic domains. In *Proceedings of the Twelfth AAAI National Conference on Artificial Intelligence, AAAI'94*, pages 1023–1028, Seattle, Washington, August 1994. AAAI Press.
- [17] Joshua I. Gold and Michael N. Shadlen. The Neural Basis of Decision Making. *Annual Review of Neuroscience*, 30:535–574, 2007.
- [18] Koosha Khalvati, Roozbeh Kiani, and Rajesh P. N. Rao. Bayesian inference with incomplete knowledge explains perceptual confidence and its deviations from accuracy. *Nature Communications*, 12(1):5704, September 2021. Number: 1 Publisher: Nature Publishing Group.
- [19] Wei Ji Ma, Konrad Paul Kording, and Daniel Goldreich. *Bayesian Models of Perception and Action: An Introduction*. MIT Press, 2023.
- [20] Benedetto De Martino, Stephen M. Fleming, Neil Garrett, and Raymond J. Dolan. Confidence in value-based choice. *Nature Neuroscience*, 16(1):105–110, January 2013. Number: 1 Publisher: Nature Publishing Group.
- [21] Annika Boldt, Charles Blundell, and Benedetto De Martino. Confidence modulates exploration and exploitation in value-based learning. *Neuroscience of Consciousness*, 2019(1):niz004, January 2019.
- [22] Rubén Moreno-Bote. Decision Confidence and Uncertainty in Diffusion Models with Partially Correlated Neuronal Integrators. *Neural Computation*, 22(7):1786–1811, July 2010.
- [23] Jeroen Brus, Helena Aebbersold, Marcus Grueschow, and Rafael Polania. Sources of confidence in value-based choice. *Nature Communications*, 12(1):7337, December 2021. Number: 1 Publisher: Nature Publishing Group.
- [24] Nathaniel D. Daw, Samuel J. Gershman, Ben Seymour, Peter Dayan, and Raymond J. Dolan. Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron*, 69(6):1204–1215, March 2011.

## NeurIPS Paper Checklist

The checklist is designed to encourage best practices for responsible machine learning research, addressing issues of reproducibility, transparency, research ethics, and societal impact. Do not remove the checklist: **The papers not including the checklist will be desk rejected.** The checklist should follow the references and precede the (optional) supplemental material. The checklist does NOT count towards the page limit.

Please read the checklist guidelines carefully for information on how to answer these questions. For each question in the checklist:

- You should answer [Yes], [No], or [NA].
- [NA] means either that the question is Not Applicable for that particular paper or the relevant information is Not Available.
- Please provide a short (1–2 sentence) justification right after your answer (even for NA).

**The checklist answers are an integral part of your paper submission.** They are visible to the reviewers, area chairs, senior area chairs, and ethics reviewers. You will be asked to also include it (after eventual revisions) with the final version of your paper, and its final version will be published with the paper.

The reviewers of your paper will be asked to use the checklist as one of the factors in their evaluation. While "[Yes]" is generally preferable to "[No]", it is perfectly acceptable to answer "[No]" provided a proper justification is given (e.g., "error bars are not reported because it would be too computationally expensive" or "we were unable to find the license for the dataset we used"). In general, answering "[No]" or "[NA]" is not grounds for rejection. While the questions are phrased in a binary way, we acknowledge that the true answer is often more nuanced, so please just use your best judgment and write a justification to elaborate. All supporting evidence can appear either in the main paper or the supplemental material, provided in appendix. If you answer [Yes] to a question, in the justification please point to the section(s) where related material for the question can be found.

IMPORTANT, please:

- **Delete this instruction block, but keep the section heading "NeurIPS paper checklist",**
- **Keep the checklist subsection headings, questions/answers and guidelines below.**
- **Do not modify the questions and only use the provided macros for your answers.**

### 1. Claims

Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

Answer: [Yes]

Justification: In the last paragraph of our introduction we outline the scope of what we are going to do in this paper and then in the following sections we complete all of the objectives that we've outlined in the introduction

Guidelines:

- The answer NA means that the abstract and introduction do not include the claims made in the paper.
- The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
- The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
- It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

### 2. Limitations

Question: Does the paper discuss the limitations of the work performed by the authors?

Answer: [Yes]

Justification: In our results section we discuss the limitations of our model through our discussion of how many subjects' behavior the model is most explanatory for compared to other Bayesian models. We also discuss limitations of scaling of the model in the background section of the paper.

Guidelines:

- The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
- The authors are encouraged to create a separate "Limitations" section in their paper.
- The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
- The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
- The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
- The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
- If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
- While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

### 3. Theory Assumptions and Proofs

Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: There are full explanations for our theoretical results and the references were provided also include detailed theoretical proofs.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

### 4. Experimental Result Reproducibility

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: We provide a lot of detail on the equations that we used to get our results and the fitting methods that we used for these equations and both of the datasets we used are freely accessible so one should have all of the information that they need to replicate results.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general, releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  - (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  - (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  - (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  - (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

## 5. Open access to data and code

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [\[Yes\]](#)

Justification: Both of the data sets that we used are freely accessible and the code that we used is attached in supplementary material and clearly commented, providing sufficient instruction to reproduce our results.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.

- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

## 6. Experimental Setting/Details

Question: Does the paper specify all the training and test details (e.g., data splits, hyper-parameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [\[Yes\]](#)

Justification: The results section is clear on what splits of the data were used for training and testing the model, which parameters were fitted, and how these parameters were fit.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

## 7. Experiment Statistical Significance

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [\[Yes\]](#)

Justification: For the main experimental results that had stochasticity in the fitting procedures for them, uncertainty bars are shown in the tables to show the variability in fitting across different runs of the model.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

## 8. Experiments Compute Resources



Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [No]

Justification: The model that we ran wasn't incredibly computationally expensive; cumulatively, all of the results could be reproduced in a couple of hours on a normal laptop so it wasn't relevant to mention in the main paper for sake of reproducibility.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

## 9. Code Of Ethics

Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines>?

Answer: [Yes]

Justification: We have viewed the NeurIPS Code of Ethics and our paper conforms to it in every respect.

Guidelines:

- The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

## 10. Broader Impacts

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [Yes]

Justification: The societal impacts of the work performed is discussed in the discussion section of the main paper.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.



- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

#### 11. Safeguards

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: Our model does not have a high risk for misuse as it is more in the realm of basic research.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

#### 12. Licenses for existing assets

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification: The original creators of the 2 open data sets that we use in are experiments are properly credited in this paper.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, [paperswithcode.com/datasets](https://paperswithcode.com/datasets) has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

#### 13. New Assets

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [Yes]

Justification: Code is included and is well documented so that other people can understand how to use it.

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

#### 14. Crowdsourcing and Research with Human Subjects

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [\[Yes\]](#)

Justification: The instructions that were communicated to human subjects and the details about compensation for these human subjects are communicated as best as we know them from the original papers where we got the data sets from.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

#### 15. Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [\[NA\]](#)

Justification: While we used data from human subjects for the main results of this paper, we didn't not conduct these experiments ourselves and so were not part of the ethical review process for these experiments. The papers we took these open data sets from didn't describe any potential harms for the study participants.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.