A Normative Comparison of Threshold Views Through Computer Simulations

Alice C.W. Huang

June 13, 2022

Abstract

The threshold view says that a person forms an outright belief P if and only if her credence for P reaches a certain threshold. Using computer simulations, I compare different versions of the threshold view to understand how they perform under time pressure in decision problems. The results illuminate the strengths and weaknesses of the various cognitive strategies in different decision contexts. A threshold view that performs well across diverse contexts is likely to be a cognitively flexible and context-dependent fusion of several of the existing theories. The results of the simulations also cast doubts on the possibility of a threshold view that is both simple enough to streamline our reasoning while also allowing us to form good action-guiding beliefs.

Keywords: credences, outright beliefs, ecological rationality, threshold view, computer simulations

1 Introduction

There seem to be two notions of belief at play in our daily discourse. On the one hand, we can ask Alexis, is Camille coming home for dinner? The answer to this question is binary. Either Alexis believes she will, or he does not. If he does, we say that Alexis *outright believes* that Camille will come home in time. Alternatively, I might ask, how sure are you that Camille is coming home for dinner? Alexis might say that he is sure, or that he is quite confident. If he wants to be precise, he might say, "well, I am 90 % sure she is going to make it." In this case, we say that Alexis has *credence* 0.9 that Camille will be home for dinner.

The tricky question is how outright beliefs relate to credences. The *threshold* view says that S outright believes P if and only if S's credence for a proposition P is higher than a given threshold t. On most, but not all, views, $\frac{1}{2} < t < 1$. The threshold view faces severe challenges by the lottery paradox (Kyburg, 1961).[1] Different variations of the threshold view have since been proposed in lieu of the original version.

This paper illustrates the strengths and weaknesses of different types of threshold views. Variations of the threshold view are categorized into four broad types based on (1) whether

[1] The lottery paradox is a problem for most threshold views with $t < 1$. Suppose there is a lottery with one winning ticket out of n tickets, where $(1 - \frac{1}{n}) > t$. For each individual ticket, S's credence that it loses suffices for an outright belief according to standard threshold views. Therefore, S outright believes that each ticket will lose. However, S also outright believes, by assumption, that at least one ticket will win. It follows that conjunctive closure is violated. In other words, if $t < 1$, then S outright believes P and Q without believing $P \wedge Q$.

credences are distributed across all possibilities, and (2) whether an agent's doxastic state is sensitive to pragmatic factors, according to the theory. I use computer simulations to compare how each of them performs in a class of time-sensitive decision problems. Their performances are assessed along the dimensions of speed and accuracy.

Outright beliefs are often thought to reduce our cognitive workload and streamline our reasoning (Friedman, 2019; Holton, 2014; Nagel, 2010). Moreover, it has been argued that one of their main functions is action-guiding (Williamson, 2000; Frankish, 2009). The threshold view has the advantage of providing a simple algorithm to translate the cognitively taxing, graded doxastic system (credences) into a manageable binary form (outright beliefs) to guide our actions.

The upshot of the simulations, however, is that simple algorithms for converting credences to outright beliefs are unlikely to perform well across a diverse range of decision problems. The relative performances of different threshold algorithms are volatile and depend on the difficulty and stakes of the tasks. A threshold algorithm that performs well uniformly is rather a cognitively flexible fusion of several of the existing theories. If we form action-guiding beliefs from credences following a simple threshold rule, we end up taking actions that lead to undesirable outcomes. On the other hand, if our doxastic strategy is flexible enough to accommodate different levels of difficulty and stakes, then the threshold algorithm becomes so complex that it no longer has the advantage of streamlining our reasoning.

I begin by presenting the motivations and variations of threshold views (section 2). In section 3, I describe the simulation and make explicit some underlying assumptions. Section 4 reports the results of the simulations and section 5 is a discussion of the implications of the results.

2 Variations of the Threshold View

2.1 Cognitive Streamlining and the Belief-Action Link

In this paper, I only focus on the threshold view. This means that I suppose two things: (a) there are outright beliefs[2], and (b) a person's outright beliefs are at least partially fixed by her credences.[3] Throughout the paper, I also endorse (c) the belief-action link (details below). Those who reject these three assumptions can still see this as a study of different decision strategies, independent of issues related to graded and binary beliefs.

There are practical reasons for forming outright beliefs in addition to credences. Compared to credences, binary beliefs are easier to keep track of for cognitively limited beings. As Holton (2014) points out:

> Unless their powers of memory and reasoning are very great, those who employ credences risk being overwhelmed by the huge mass of uncertainty that the approach generates. First they will have to store very much more information:

[2]Justifications for this assumption can be found in our linguistic practices (Maher, 1993), as well as from the perspective of cognitive efficiency, which I explain below. Moreover, see Weisberg (2020) for arguments from the perspective of psyontology that we have both binary and graded beliefs.

[3]That outright beliefs are at least partially fixed by credences is the core idea of the threshold view. Alternatives to the threshold view include, first, theories according to which graded and binary beliefs are entirely independent. This leads to *prima facie* strange results, for in virtue of what do the two systems both count as belief states if there is no connection between them? Secondly, graded beliefs might be fixed by binary beliefs. This approach is favored by Holton (2014) and Easwaran (2016), who think that graded beliefs are reducible to binary beliefs.

> rather than just discarding the propositions that aren't believed and focusing on those that are, they will have to keep track of all of them and their associated credences. Second, they will have to be able to deploy the complicated methods needed for probabilistic reasoning. (p. 14)

In addition to the constraints on storage capacity and computational power, our time is limited. Some have argued that outright beliefs can be considered a kind of "settled opinion" that marks the end of inquiries (Friedman, 2019; Nagel, 2010). Friedman (2019), for instance, argues that "belief has some inquiry-theoretic properties that even very high and maybe even maximally high credence don't have. In particular, believing while inquiring is a form of incoherence but having even very high credence while inquiring is not." Binary beliefs enable agents to close off inquiries within a time limit.

Closing off inquiries by forming outright beliefs also streamlines our practical deliberation. In *Knowledge and Its Limits*, Williamson (2000) says:

> What is the difference between believing P outright and assigning P a high subjective probability? Intuitively, one believes P outright when one is willing to use P as a premise in practical reasoning. Thus one may assign P a high subjective probability without believing P outright, if the corresponding premise in one's practical reasoning is just that P is highly probable on one's evidence, not P itself. (p. 99)

Frankish (2009) holds a similar view and adds that "in adopting premises and goals we commit ourselves to performing any actions they dictate." We often use outright beliefs as premises when reasoning about what to do, since, given our cognitive and time constraints, it is generally more feasible than reasoning purely with credences and keeping track of even the slightest uncertainty. The belief-action link connects outright beliefs to our premise policy and actions, offering an important explanation to why binary beliefs are essential to limited beings like us.[4]

Taking stock, given that humans have limited memory storage, computational power and time, it is beneficial for our actions to be guided by outright beliefs. There may be other purposes to outright beliefs. However, since the paper aims to evaluate the performance of different threshold views in time-sensitive decision contexts, I focus on the action-guiding role of beliefs. The threshold view, if successful, offers a simple algorithm for us to read off an agent's credences (and pragmatic circumstances etc. on some variations, to be explained below) to infer the beliefs guiding their actions.

2.2 Carving up the Conceptual Space

To evaluate which formulations of the threshold view allow us to form good action-guiding beliefs in time-sentsitive decision problems, two questions are particularly pertinent. We may carve up the conceptual space within the family of threshold views according to their answers to these questions.

1. Do pragmatic factors, such as stakes, partially determine what an agent believes?

2. Are credences distributed across all the possible alternatives?

[4] Similar principles establishing the link between knowledge and action are endorsed in Fantl and McGrath (2002), Hawthorne (2003); Hawthorne and Stanley (2008), and Ross and Schroeder (2014).

These two questions are directly relevant for the assessment of threshold views in decision contexts with realistic constraints. Doxastic sensitivity to pragmatic factors optimizes the allocation of cognitive resources, enabling us to dedicate more time and efforts to high-stake cases. Restricting the domain of credence distribution is a way to economize by focusing only on the most germane hypotheses. Furthermore, pragmatic sensitivity and the scope of credence distribution can be modelled computationally for empirical assessment, providing a good starting point to evaluate the action-guiding function of outright beliefs formed according to various threshold views.

Many have observed that pragmatic factors such as stakes seem to play a role in determining what an agent believes (Stanley, 2005; Fantl and McGrath, 2002; Hawthorne, 2003; Ross and Schroeder, 2014). For instance, in normal circumstances, I believe outright that the earliest metro in Toronto departs at 6 a.m. Call this proposition M. I go about my daily life treating M as a true proposition and act accordingly. Now suppose that I have a flight from Toronto to Paris on the next day at 8 a.m. Even though there is no reason for me to think that the timetable has changed, I begin to check the Toronto Transit Company website to confirm that the metro starts running at 6 a.m. I want to make sure that I can take the metro to the airport and not miss my flight. My need to confirm M indicates that I no longer outright believe M, given the belief-action link. Since there is no reason for me to think that the timetable has changed, presumably my credence for M remains the same even though I ceased to outright believe M. This can be explained by the difference in the stakes involved. In normal circumstances, being wrong about M has little consequence, while on the morning of my flight, being wrong about M comes at a high cost. Doxastic sensitivity to pragmatic factors accommodates observations like this one—when stakes are high, we are disinclined to form outright beliefs.

Not all agree, however, that doxastic sensitivity to pragmatic factors should play a role in our theory of beliefs. For instance, Greco (2013) and Rubin (2015) have both proven that agents whose credences are pragmatically sensitive are irrational, in the sense that they are subjected to diachronic Dutch-books.[5] Some variations of pragmatically sensitive threshold views, notably those defended by Clarke (2013), Greco (2015) and Gao (2019), will therefore encode this kind of irrationality.[6]

The two opposing camps divide the family of threshold views into two categories—variations which account for pragmatic sensitivity and variations which do not.

The second question on the scope of credence distribution is motivated by concerns that traditional Bayesian epistemology is too idealized to be applicable to human agents. The ideal Bayesian agent keeps track of all possibilities and constantly updates them via conditionalization. Since humans are limited in cognitive capacities, some have argued that a fully general domain over which our probabilistic credences are distributed is untenable.

Holton (2014) for instance says that "for limited creatures like us, reducing the options to a number of fixed points is essential. It...provides a way to keep the time spent on deliberation in check." Greco (2015) likewise thinks that on a realistic model of human doxastic states, "we don't have fully domain-general representations that guide our actions no matter what our situation. Rather, whatever task we're engaged in—with 'task' understood broadly, so that 'tasks' include conversation, and even just internal deliberation—we'll only treat certain possibilities as live, and which possibilities we do treat as live will depend on a wide range of

[5] Greco (2015) later turns to support what I will call "SmallSens(1), " or "credences-one sensitivism" and accept irrationality as an integral part of beliefs.

[6] Being subjected to a diachronic Dutch-Book might not be a mark of irrationality in some views. See for example Christensen (1991).

factors, including broadly 'practical' ones." Clarke (2013) also argues that credences should only be assigned to a context-depend subset of *serious possibilities*. In addition, Norby (2015) draws from the psychology literature to show that credences are assigned "on the spot" to possibilities determined to be *live* according to various psychological mechanisms. This ensures that we only deal with a subset of possibilities tractable for the working memory.[7]

I call possibilities to which credences are assigned *active possibilities* to remain neutral on different understandings of "live" or "serious" possibilities. For the purpose of this paper it is enough to proceed with just a rough idea of what they are, without a precise definition.

The family of threshold views can again be categorized into two kinds depending on whether credences are assigned to a full general domain, or only the active possibilities. Together, the two orthogonal questions carve up the conceptual space into four categories. Table 1 shows the four types and how they will be labeled in the remainder of this paper.

	Epistemic Purist	Pragmatic Sensitivist
Full Domain	**FullPure**	**FullSens**
Small Domain	**SmallPure**	**SmallSens**

Table 1: Four types of threshold views and how they will be short-handed in the remainder of this paper. Epistemic purism does not incorporate doxastic sensitivity to pragmatic factors, whereas pragmatic sensitivism does.

Each of these four types of threshold view provides a simple algorithm that allows us to infer which outright beliefs an agent holds based on their credences and what is at stake. I argue, however, that no threshold view simple enough to fit into these categories can produce good action-guiding beliefs across a diverse range of realistic decision problems. Below I introduce the four types of threshold views in more details.

2.3 Full Domain Purism

Out of the four types of threshold view, full domain purism is the most idealized. What an agent believes is fully fixed by truth-relevant factors, and when the agent's credence for P reaches a fixed threshold, she outright believes P. The agent's credences are distributed across all possibilities.

This category consists notably of the original threshold view, often attributed to Locke and termed the "Lockean thesis" (Foley, 1992).

2.4 Full Domain Sensitivism

Full domain sensitivism takes into account pragmatic sensitivity, but does not make the distinction between active and inactive possibilities. Credences are distributed across all possibilities.

In this paper, I consider stakes to be the only kind of pragmatic factor for the sake of simplicity, though other practical parameters also influence our doxastic states. In the

[7] Weatherson (2005) has a similar spirit of anti-idealization, though he focuses on limiting the domain where conjunctive closure should be obeyed, as opposed to limiting the set of active possibilities. He argues that we can have probabilistic coherence without logical coherence, and in particular, that only the set of salient propositions must be conjunctively closed. See also Weisberg (2020) for more on how credence is assigned on the fly, and Stanley (2007) for a normative construal of active possibilities.

simulations, it is implicitly assumed that all the other pragmatic factors are fixed across different contexts. The relevant notion of stakes here is perceived stakes. It is how high the agent perceives the stakes to be, as opposed to how high the stakes actually are.

The pragmatic sensitivity of agents can be the result of two different mechanisms. First, the threshold of belief may vary across contexts. In high-stake scenarios, the threshold is higher, so a higher level of confidence is required for outright beliefs. In low-stake cases, the threshold is lower, and the agent more easily acts as though an uncertain proposition is true. In his influential paper on pragmatic encroachment, Weatherson (2005) says that "interests matter not because they affect the degree of confidence that an agent can reasonably have in a proposition's truth. (That is, not because they matter to epistemology.) Rather, interests matter because they affect whether those reasonable degrees of confidence amount to belief. (That is, because they matter to philosophy of mind.)" Ganson (2008) likewise contends that "in order to count as believing P in a range of circumstances, one must be willing to act as if P in those circumstances: one's degree of belief that P has to be high enough that one is willing to act as if P under those circumstances."

Gao (2019) argues, however, that intuitive and psychological evidence indicates otherwise.[8] The threshold does not change across contexts. Instead, the agent's credences change directly as a matter of contingent, psychological fact. What we outright believe varies as a direct result of the change in credences (Gao, forthcoming).

As we will see later, the implementations of these two branches of full domain sensitivism are the same under the assumptions in the simulations and can therefore be interpreted both ways.

2.5 Small Domain Purism

According to small domain purism, the threshold is fixed and credences are only distributed over a subset of active possibilities. Since there is no pragmatic sensitivity for small domain purism, whether a possibility is active cannot be fixed by stakes (as is the case for small domain senstivism below). I implement the selection of active possibilities by imposing a fixed size for the domain that is constant across different decision problems. In the simulations, each small-domain agent updates and keeps track of 4 active possibilities that are the most probable given their prior probabilities. (Details in the next section.)

The number 4 is informed by studies in psychology on working memory. Working memory is the part of memory used to plan and carry out behaviors, manipulate information and perform cognitive tasks that require attention (George A. Miller et al., 1960; Cowan, 2008). It has been argued that the capacity of our working memory is around 4 items. (See Cowan et al. (2008) for an extensive review.)

Claims about the "magic number" 4 are not uncontroversial and especially complicated when applied to possibilities. Firstly, in the psychology experiments conducted, the items that participants are asked to memorize are usually chunks of visual or verbal cues, and sometimes a mix of both, but it is not obvious how a possibility or a hypothesis can be regarded as a "chunk" or an item. Secondly, the capacity of our working memory may alternatively be understood in terms of resources, in which case its limits should be measured on a continuous, as opposed to discrete, scale. It is not clear then how we can constrain the range and size of active hypotheses. Lastly, the capacity of our working memory is not the only relevant factor that determines how many hypotheses we are cognitively capable

[8] See Mayseless and Kruglanski (1987) for the psychological perspective on this topic.

of keeping track of, so simply importing the standard for working memory is inevitably unsatisfactory.

Much is left to be explored in terms of how to realistically model the domain of active possibilities. I tentatively choose 4 in the simulations based on the best currently available evidence.

Despite some degree of arbitrariness, this implementation captures the main features of active possibilities. Cowan (2010) contends that the limit in capacity may not be only a weakness, but also a strength:

> Mathematical simulations suggest that, under certain simple assumptions, searches through information are most efficient when the groups to be searched include about 3.5 items on average... A relatively small central working memory may allow all concurrently active concepts to become associated with one another (chunked) without causing confusion or distraction. Imperfect rules, such as those of grammar, can be learned without too much worry about exceptions to the rule, as these are often lost from our limited working memory. (p. 56)

This analysis resonates well with the motivation for a smaller domain, as well as some results of the simulations that we will see later. Less is sometimes more. The ability to ignore some possibilities might be an advantage to the overall efficiency of decision making in practical, realistic scenarios. Furthermore, this implementation captures the idea that agents reduce their cognitive workload by ignoring *remote* possibilities, since the four active hypotheses are the most probable ones given their prior information.

To my knowledge, no one has proposed small domain purism as a model for beliefs, but I include it in the simulations for the completeness of the conceptual space.[9]

2.6 Small Domain Sensitivism

Small domain sensitivism combines active domains with doxastic sensitivity to pragmatic factors—which possibilities are active is partially determined by the agent's practical circumstances. The higher the stakes, the more cognitive resources an agent allocates to a given task. As a result, the agent takes into account more remote possibilities for the sake of prudence. By contrast, when the stakes are low, the agent's active domain is smaller.

There are many possible variations of small domain sensitivism. The credence-one view holds that the threshold for outright beliefs is always 1. To outright believe a proposition, subjective certainty *given the subset of active possibilities* is required. The practical context only *indirectly* influences what the agent believes by influencing what possibilities are active (Clarke, 2013; Greco, 2015). This is different from standard threshold views, which usually assume the threshold to be higher than 0.5 and lower than 1. The credence-one formulation of small domain sensitivism is, to my knowledge, the only form of small domain sensitivism that has been proposed in the literature.

However, it turns out that the credence-one strategy performs far worse than any other view across every decision context. In view of this result, it is more interesting to also consider a different version of small domain sensitivism. There are plenty of options. (*i*) The threshold of belief might be a constant $\frac{1}{2} < t < 1$ while the active domain varies depending on stakes, (*ii*) practical interests might influence both the threshold and the domain size, or (*iii*) the active domain might be constant, and stakes affect only the threshold. These other

[9] Babic (2019) endorses a pragmatic view similar to small domain purism, though he does not address the belief-credence connection, nor does he explicitly mention the concept of an active domain.

options have not been considered in the literature as far as I am aware. For simplicity, I include the first of the three options, the formulation combining a pragmatically sensitive small domain with a fixed threshold below 1, and leave the others for future work.

3 The Simulation

3.1 The Problem

Most problems we face require not only accurate beliefs, but also decisions made in a timely manner. Should I book a vacation home early for a cheaper price or later to compare more options and dates? Should I make a witty comment now or think twice to avoid offending anyone but risk missing the moment for a laugh? Taking time to think always comes at a cost, even when the cost is not obvious. More often than not, the time spent on deliberating is time not spent on acting, and at some point one has to close off the inquiry and take action. This leads to a strategic question—when is a good time to stop thinking and start acting? How many career options should I explore before settling on philosophy? How certain do I need to be that my partner is *the one* before I get married? Since we are almost never able to reach absolute certainty given the limited time and energy we allocate to each inquiry, choosing the appropriate stopping point is important to the success of our reasoning. We usually need to act without being certain.[10] Moreover, in reality, we often face multiple different decisions at once. With limits on our time and cognitive energy, prolonging deliberation on one issue means taking away time available for other issues.

The trade-off between making up our minds quickly and forming accurate beliefs is a kind of *speed-accuracy trade-off*.[11] The speed-accuracy trade-off is an extensively studied phenomenon in cognitive science. To a first approximation, it is assumed that the more time an agent spends on collecting information, the more accurate her beliefs will be. Agents attempt to strike a good balance between speed and accuracy, only spending more time on a given deliberation if she judges that the reward for increased accuracy is worth the extra efforts.

The structure of the problems that agents face in the simulations can be characterized by three main features: (1) forming an accurate outright belief leads to a higher score than forming none, (2) forming an inaccurate outright belief leads to a lower score than forming none, and moreover, (3) the longer it takes for an agent to form an outright belief, the lower their score. This setup allows us to evaluate the cognitive performance of different types of agent by how well they balance speed and accuracy given the decision context.

In each run, each agent is given at most 100 time steps to solve a problem, and each receives a score at the end. The agents can be thought of as doctors who, under time pressure, need to find out which disease their patients have and intervene accordingly. The patients' probability of death is modelled by a cumulative density function of some Weibull distribution, a common and flexible parametric family for modeling time to event data, such as survival (appendix A). The probability of a patient dying without any intervention increases with each time step.[12]

[10] See Nagel (2010) for an extensive discussion about the termination of inquiries.

[11] See Duckworth et al. (2018) 2.2.1 for a brief introduction.

[12] The Weibull distribution is determined by two parameters—shape and scale. The shape parameter is randomized between 1 and 2, and the scale is fixed at 50. The two parameters together determine, roughly, how rapidly the patient's probability of death increases with each time step. For instance, when the shape parameter is small, the probability of death increases faster in the beginning and slower towards the end.

Doctors intervene once an outright belief is formed. For simplicity, it is assumed that each possible disease has its unique treatment. The effects of interventions are determined by two parameters which specify how much a correct intervention increases the patient's chance of survival and how much an incorrect intervention increases a patient's chance of death, given their baseline chance of survival at each time step.[13] Once an intervention is made, the simulation terminates and the score is calculated. If no intervention is made, i.e., no outright belief is formed, then the simulation terminates after the 100^{th} time step.

In each run, for each doctor, the curve of their patient's probability of death and the effects of interventions vary, as the parameters are randomized. With 3000 runs conducted for each type of threshold agent, this ensures that the results of the simulations are robust across different problems of the same structure, and not mere artifacts of specific parameter settings. Fig. 1 shows two examples of patients with different parameters.

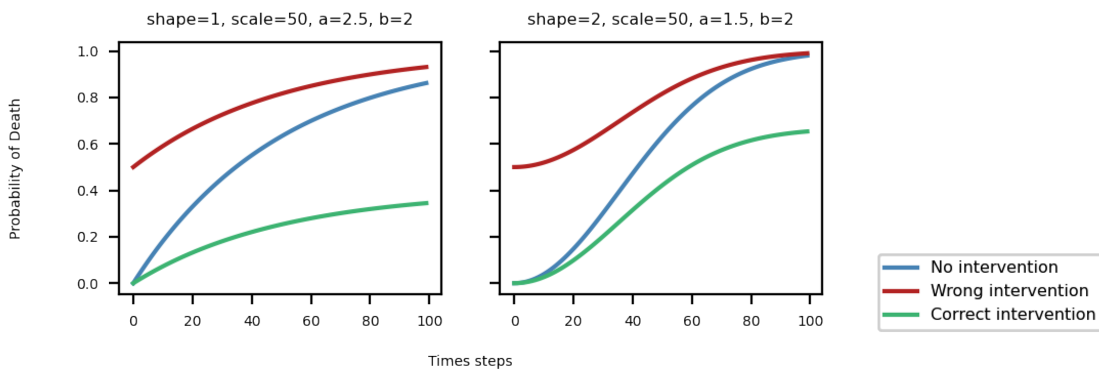


Figure 1: Examples of patients' probability of death in 100 time steps. Parameters *shape* and *scale* determine the blue line in the graph, which is the baseline probability of death without intervention. Parameters a and b determine the effects of interventions.

In each run, the doctors must find the true hypothesis amongst 5 to 15 jointly exhaustive and mutually exclusive hypotheses. Each hypothesis is represented by a unique real number between 0 and 1, and can be understood as a disease with a unique treatment or intervention. I assume that agents know what the possible alternatives are. In the medical diagnosis scenario, this amounts to saying the doctors having exhaustive knowledge about what diseases could possibly cause the symptoms that the patients experience. In reality, this tends not to be the case. There are often possibilities that one might miss, and finding the complete and correct set of possibilities is a whole other cognitive endeavor. In order for the simulations to remain tractable, I set aside issues related to identifying the hypothesis space.

At every time step, the agents receive a new piece of evidence and update their credences for each hypothesis following the standard Bayesian formalism. The initial probability for

The smaller the scale is, the faster the patients' probability of death increases with time. My setup of the problem here roughly follows that of Douven (2020).

[13] The effects of correct and incorrect interventions are determined by parameters a and b. Let the patient's baseline probability of death without intervention at a given time be p. The correct intervention brings down the probability of death to p/a, whereas a wrong intervention increases it to $\frac{(p+1)}{b}$. The larger a is, the more effectively a correct intervention lowers the patient's risk of death. The smaller b is, the more a wrong intervention increases the patient's probability of death. a is randomized between 1.5 and 2.5 whereas b is fixed at 2.

each hypothesis is determined by a random sample from a symmetric dirichlet distribution, which is often used as prior distributions for multivariate discrete data (appendix B).[14] The priors used in the simulations are more realistic than uniform priors, since we rarely begin an inquiry with no preconception of what the underlying probability distribution is like. Instead, we treat some hypotheses as more likely than others from the get go.

In the simulations, it is assumed that all agents have initially accurate prior probabilities, though in reality our prior probabilities are often inaccurate. Following the diagnostic narrative, this amounts to saying that all doctors know how common each disease is. The disease in each medical case is randomly chosen based on the distribution specified by each doctor's prior probabilities. Once the disease of the patient is determined, the agent tries to find out which one it is with the help of test results.

The hypotheses are represented by the bias of a coin and evidence or test results are represented by the outcome of coin flips. Although coin biases are represented by real numbers, the relevant hypotheses (interventions) here are categorical. Therefore, a miss is as good as a mile.[15] At each time step the coin is flipped once, and the agents update their credences based on the outcome. In reality, evidence does not always come in the form of sampling data. For example, we also obtain information indirectly from testimony and need to assess the reliability of the source. Modelling the collection of data as flipping biased coins sets aside some of the complexity of real life evidence gathering.

At each time step, the doctor has $\mathrm{N}_{hyp} + 1$ choices, with N_{hyp} being the total number of mutually exclusive and jointly exhaustive possibilities in the problem. She can choose any of the N_{hyp} interventions, or choose to wait and collect more data.

3.2 Stakes

In each run, every doctor is paired with a diagnostic case. Each case has a stake parameter, which is a randomly-drawn integer between 1 and 5 (inclusive). The stake parameter represents the number of patients in a cluster infection. Stake 1 means that only one patient's life is on the line, whereas stake 5 means that a family of 5 have the same disease and the same symptoms.

If an agent intervenes before 100 time steps, she scores the sum of all the patients' probability of survival. In other words, the score an agent receives is calculated by how many patients are expected to survive. For instance, suppose the doctor has a cluster infection of 3 patients. At time t she intervenes, and given her intervention, the probability of death for the patients becomes p. In this case, she receives $(1 - p) \times 3$ points. [16]

[14] The symmetric dirichlet distribution is determined by a concentration parameter. The larger the concentration parameter is, the more equally probability is distributed across the hypotheses. By contrast, when the concentration parameter is small, the agent has high initial credences for a small number of possibilities and low credences for the other hypotheses. In the simulation, the concentration parameter is randomly chosen from a normal distribution centered at 0.5, with a standard deviation of 0.1. The reason for restricting the concentration parameter to a limited range is to prevent having priors too close to the uniform prior or too concentrated and thereby avoid giving small domain agents unfair advantage or disadvantage.

[15] This simulation can thus also be interpreted simply as a guessing game of coin biases with monetary rewards involved.

[16] Here I provide a more detailed example of the scoring rule. Suppose again that the doctor has a cluster infection of 3 patients, and the probability of death for each patient in the case at time t is 0.6. If the agent's diagnosis is correct, she gains $(1 - \frac{0.6}{a}) \times 3$ points. If the intervention is incorrect, the agent gains $(1 - \frac{1.6}{b}) \times 3$ points, with a and b being the parameters explained in footnote 13. If after 100 time steps, no intervention has been made, then the agent gets $(1 - p_{100}) \times 3$ points, where p_{100} is the patient's probability of death at the 100^{th} time step.

Sensitivist agents adjust how their credences translate to beliefs based on whether it is a low- or high-stake case, whereas purist agents do not. The numbers that represent stakes are chosen arbitrarily, but the exact numbers do not matter, since we will focus on analyzing ordinal statistics, as opposed to the absolute scores that doctors receive. As we will see shortly, sensitivist agents make adjustments based on stakes only after the numbers are re-scaled to $[-1, 1]$ in order to remove the effects of the arbitrarily chosen numbers representing stakes.

Stakes are assumed to be known to the agents. In reality, this is rarely the case. My decision to turn left on an evening walk might lead me to a rendezvous with the love of my life or to a dangerous robbery. Unaware of what will happen, I make my decision without taking these "real" stakes into account. Finding out the hidden costs and rewards is nearly impossible and cognitively costly. In this paper, I set aside issues related to hidden stakes and take the stakes simply to be fully known to the agent.

3.3 Four Types of Agent

The four types of threshold views are modelled as follows. FullPure agents form the outright belief that a certain hypothesis is true as soon as their credence for that hypothesis is above 0.9. As soon as the agent outright believes a hypothesis, she intervenes.

The threshold for FullSens varies with stake. Each agent has a sensitivity parameter ϕ, which is a real number randomly chosen from the normal distribution $\mathcal{N}(0.5, 0.1)$.[17] With sensitivity 1, subjective certainty is required for outright belief in cases with the highest possible stake (5). An agent with sensitivity 0 is the same as a FullPure agent. For every unit change in stake, an agent with higher sensitivity changes her threshold more than an agent with lower sensitivity. The function $f : \mathbb{Z} \to [-1, 1]$ re-scales the stake of a case such that the lowest stake is mapped to -1 and the highest stake is mapped to 1.[18] Finally, the threshold of a FullSens agent is partially fixed by the default threshold τ. The default threshold is set to 0.9 in order to be consistent with other types of agent. Let s denote stakes. The stake-sensitive threshold of a FullSens agent is given by the following function:

$$T(s, \tau, \phi) = \tau + (1 - \tau) \cdot \phi \cdot f(s) \tag{1}$$

For example, a FullSens agent with sensitivity $\phi = 0.8$ will have a threshold of $0.9 + 0.1 \cdot 0.8 \cdot f(5) = 0.98$ for cases where the stakes are the highest, requiring nearly subjective certainty to outright believe that a hypothesis is true. With the lowest stake, the same agent has a threshold of only $0.9 + 0.1 \cdot 0.8 \cdot f(1) = 0.82$. A FullSens agent with a lower sensitivity $\phi = 0.5$ will have a threshold of $0.9 + 0.1 \cdot 0.5 \cdot f(5) = 0.95$ for $s = 5$ and $0.9 + 0.1 \cdot 0.8 \cdot f(1) = 0.85$ for $s = 1$.[19]

This implementation can be interpreted as either shifting threshold or shifting credences directly. It is easy to see why the current implementation can be understood as shifting threshold. On the other hand, if we want to model the version of FullSens that Gao (2019) advocates, according to which agents shift credences directly in response to changes in stakes, credences for different hypotheses will be systematically higher/lower for low-/high-stake cases while the threshold remains the same. Whether it is credence or threshold that is

[17] The normal distribution is truncated at 0 and 1.

[18] In the simulations, since the stakes range from 1 to 5, $f(s) = \frac{s-3}{2}$. This yields the desired transformation with $f(1) = -1$ and $f(5) = 1$.

[19] When $s = 3$, the threshold is the default τ regardless of sensitivity, since $f(3) = 0$, as specified in footnote 18.

shifted, agents will end up forming the same outright beliefs at the same time steps, so the same implementation can accommodate both interpretations.

SmallPure agents have a threshold of 0.9 regardless of stakes. After the initial prior distribution is randomly generated and the true intervention is determined accordingly, SmallPure agents contract their domain to include only four active possibilities. The four most probable hypotheses are kept but the rest are discarded. Her credences are then renormalized. There is a chance that small domain agents discard the truth from the outset.

Finally, I implement two versions of SmallSens, starting with the credence-one version, denoted by "SmallSens(1)." Since credence-one agents turn out extremely maladaptive compared to other types of agent in the kind of decision problems I model, I also consider the version of SmallSens that combines a sensitivist small domain with a fixed decision threshold below 1. I denote this by "SmallSens(< 1)."

The size of the domain for a SmallSens agent is within the range of 4 ± 2, with 2 and 6 achieved at the lowest and highest stakes respectively. The higher the stakes, the more hypotheses a SmallSens agent keeps track of.[20] A SmallSens(1) agent forms an outright belief and intervenes only when all but one hypothesis are ruled out. A SmallSens(< 1) agent forms an outright belief when her credence is above 0.9.

The idea behind the functions for adjusting thresholds and domain sizes is that high-stake cases warrant more prudence. Therefore, the agent devotes more time and cognitive resources to the problem at hand and a higher level of confidence is required before outright believing a hypothesis.

4 Results

Data was collected from 3000 runs for each agent type and each of the 11 levels of difficulty ($\mathrm{N}_{hyp} \in 5...15$).[21] Statistical analysis shows that each additional possible hypothesis in the problem lowers the average score of agents by roughly 0.1 points in each run. For the sake of brevity, below I will simply call problems with fewer hypotheses *easy problems* and those with more hypotheses *difficult problems*, even though the level difficulty can result from many other factors.

Fig. 2 is a summary of the score distribution for different types of agent, averaged across all levels of difficulty and stakes. The scores reflect their ability to strike a balance between speed and accuracy, since both inaccuracy and delayed decision making are penalized. Fig. 2 demonstrates the obvious inferiority of SmallSens(1) agents. In fact, they perform significantly worse than all other agents in every combination of stake and difficulty. This suggests that forming outright beliefs only when all the other active possibilities are ruled out is maladaptive when faced with time-sensitive decision problems.

The poor performance of SmallSens(1) agents is a reflection of their inability to find a good speed-accuracy balance. Their active domain has to be small enough (and they must be lucky enough) for credence one to be possible. However, when the domain is small enough, they risk ruling out the truth from the outset. In 75% of the trials, SmallSens(1) agents are unable to form an outright belief before the simulation terminates after 100 time steps. In 97% of the cases where the stake is 5, SmallSens(1) agents form no outright belief. While in low-stake cases (stake = 1), they take action within 100 time steps in 63% of the trials, 38%

[20] Adjustments are made for cases where the domain size may be larger than N_{hyp}.

[21] The code for the simulations can be found in appendix C.

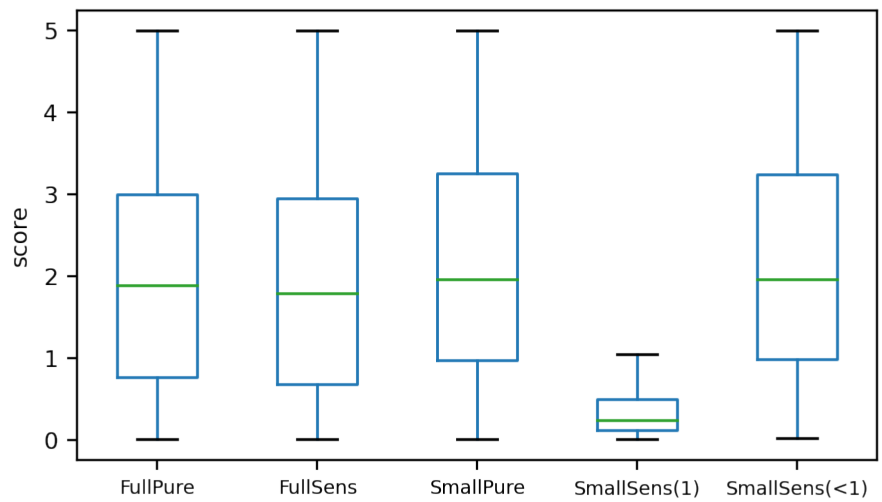


Figure 2: An overview of the average scores.

of the time the truth is not in the active domain to begin with, since their active domain includes so few hypotheses.[22]

This does not necessarily mean SmallSens(1) is a failure *tout court*. When evidence is acquired by sampling, as is the case here, it is hard to completely rule out all but one possibility, but other types of evidence may enable agents to rule out possibilities more easily. Proponents of SmallSens(1) may also introduce some other mechanisms, such as salience, to explain how we can more easily ignore uncertainty to obtain credence one. The result here nevertheless raises some worries for SmallSens(1), echoing the concern Gao (forthcoming) has—even within a small active domain, it is still too difficult to have absolute certainty. SmallSens(1) is still too close to skepticism. In what follows, I will set aside SmallSens(1) and focus on SmallSens(< 1) instead.

In fig. 2, other than SmallSens(1), the four other types of agent perform more or less equally well, when the scores are aggregated across different problem settings. However, when we look at different levels of difficulty and stakes separately, there is significant variability between different threshold strategies—they fare well in different decision contexts. Therefore, an optimal belief-forming strategy is one with cognitive flexibility and not any single variation of the threshold view that has thus far been proposed. Fig. 3 shows the average scores broken down by different decision contexts. The graph displays only the extremities, including combinations between the highest/lowest stakes, and the easiest/hardest problems. The change in performance across different decision contexts should nonetheless be understood as a gradual transition, with intermediate levels of difficulty and stakes in between.[23]

There are a few immediately noticeable observations:

[22]Situations in which SmallSens(1) agents do form outright beliefs and take actions include (a) when the true coin bias is either 0 or 1, (b) when only one hypothesis in the active domain is neither 0 nor 1 (c) their credences are automatically rounded up to 1 by the computer, which happens at around the 18^{th} decimal place.

[23]The bars are generated from data averaged over runs with $N_{hyp} = 5, 6$ for easy problems and over runs with $N_{hyp} = 14, 15$ for difficult problems.

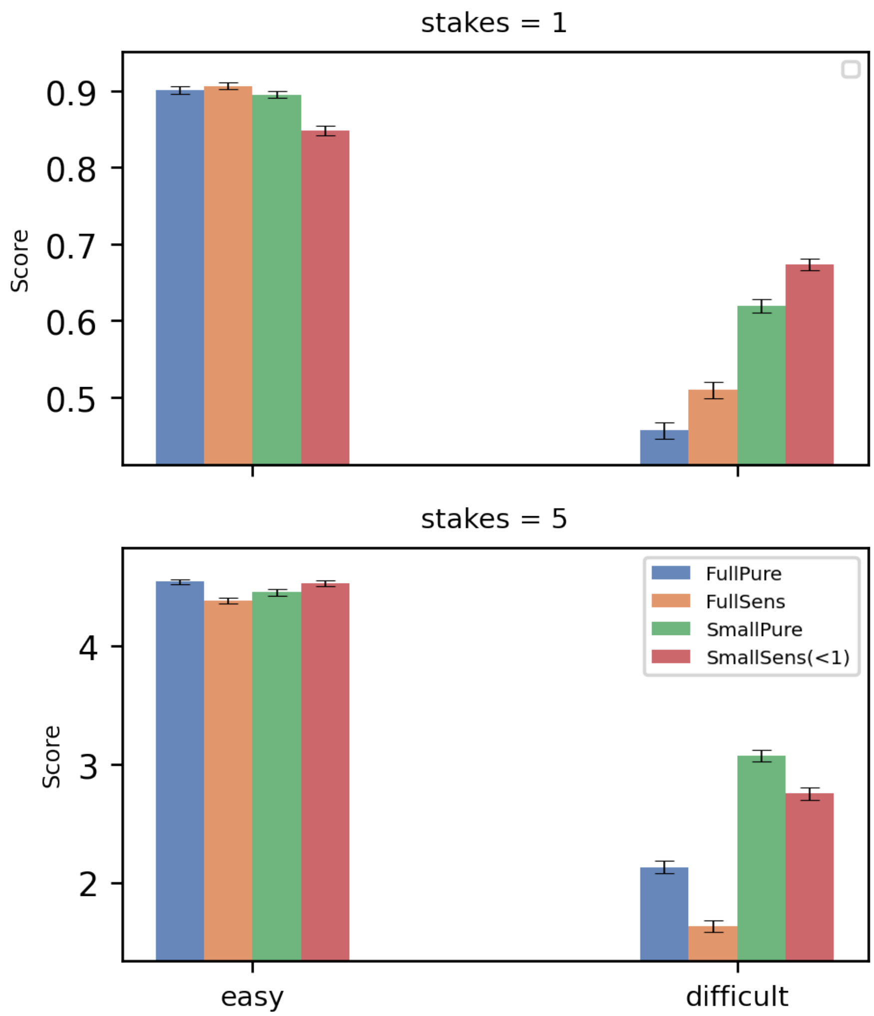


Figure 3: An overview of the average scores for different types of agent in different decision contexts. Note that the magnitude of the scores in low-stake problems is smaller than in high-stake problems. Since the scores reflect the expected number of survivals, the higher the stakes (number of patients) are, the higher the scores will be, regardless of the courses of action taken. Moreover, as expected, the average scores across all types are higher for easy problems than difficult ones.

1. In easy problems, the difference in performance between full and small domain strategies is small. In particular, in low-stake cases, full domain agents have some advantage.

2. In difficult problems, small domain agents have considerable advantage.

3. In difficult problems, sensitivist agents have an advantage in low-stake cases compared to their counterparts with similar domain sizes, but a comparative disadvantage in high-stake cases.

We can explain these results by breaking down the two dimensions of cognitive performance—speed and accuracy—that the scoring rule is designed to capture.

The first two observations can be explained together. Since all agents are able to terminate their inquiries within a short time when the hypothesis space is small, the advantage of faster decision making is not decisive in easy problems. The advantage of higher accuracy offsets the disadvantage of longer deliberation time. On the other hand, as the number of possible hypotheses increases, speed becomes very decisive and small domain strategies stand out.

To test this theory, we can analyze speed and accuracy in different decision problems. The left panel of fig. 4 shows the average number of time steps it took agents to form outright beliefs. In easy problems, the difference in speed is small, whereas the discrepancy is large in difficult problems.[24] The right panel shows the average accuracy score of outright beliefs (if any). An accuracy value of 1 or -1 is assigned to correct and incorrect intervention. If no diagnosis is made within 100 time steps, then 0 is assigned.

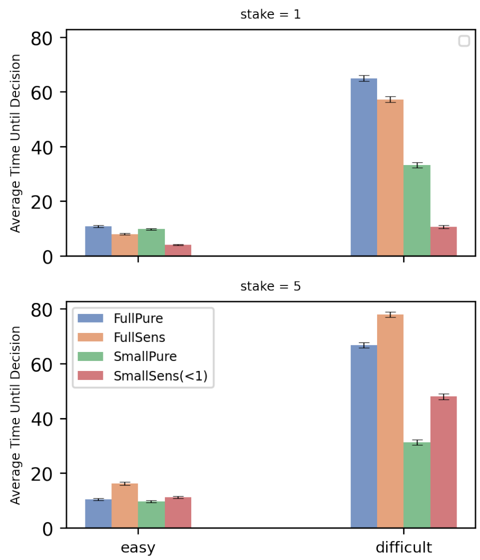
Comparing fig. 3 and fig. 4, the relative accuracy between different types of agent (except for FullSens) roughly predicts relative scores in easy problems, whereas speed does not. In fact, faster speed roughly correlates with lower overall score (again, with the exception of FullSens, to be addressed below).

On the other hand, in difficult problems, speed clearly predicts high score, whereas accuracy does not. In particular, SmallSens(< 1) agents have the lowest accuracy for low-stake difficult problems, but still come out on top in terms of overall performance due to their huge advantage in speed.

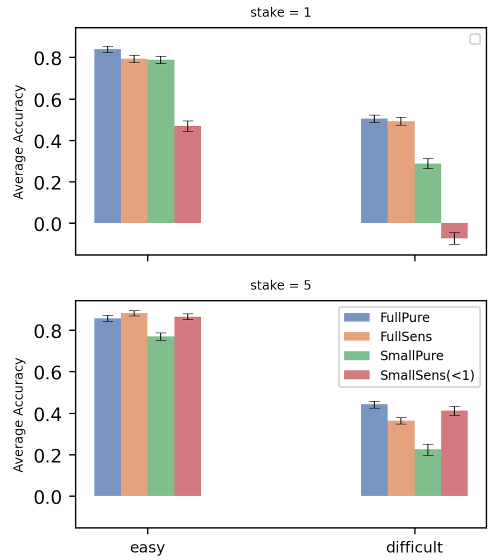
The third observation can likewise be explained from the perspective of speed and accuracy. Sensitivist agents become more prudent in order to gain accuracy when a lot is at stake. They spend more time collecting data before making up their minds. In the present simulation, prudence slows down FullSens agents so much that it becomes a comparative disadvantage when the problem is difficult. Comparing the upper subplots (low-stake) to the lower (high-stake) in fig. 4, the deliberation time for sensitivist agents is indeed longer when the stakes are high. However, while increased deliberation successfully increases accuracy for SmallSens(< 1) agents, it decreases accuracy for FullSens agents. There is a failure of speed-accuracy trade-off. Without reducing the size of the domain, increased thresholds in high-stake problems render FullSens incapable of making up their minds.

Fig. 5 shows the relation between stake and accuracy for different types of agent. The left panel shows data from all trials, including those with no intervention. As expected, there is no relation between accuracy and stake for purists. What is unexpected is that for FullSens, there is in fact an overall negative correlation between stake and accuracy, contrary to the motivation for sensitivism. The right panel shows only trials where an outright belief

[24] A related observation not shown in this graph is that the performances of small domain strategies are less affected by difficulty than full domain strategies. However, since the more possible alternatives there are, the more likely the truth is outside of the most probable 4 ± 2 hypotheses, there is still a negative correlation between the average scores and difficulty for small domain agents.



(a) Speed



(b) Accuracy

Figure 4: (a) The average number of time steps taken before forming outright beliefs in different decision contexts. (b) The average accuracy of interventions where 1 corresponds to correct intervention, 0 to no intervention and -1 to incorrect intervention.

is formed by the end of 100 time steps. Among these, stake positively predicts accuracy for all sensitivist agents. This shows that the inferiority of FullSens in high-stake cases is a result of slow or no decisions, exacerbated by a large domain. Prudence pays off, but only when a decision can be made.[25]

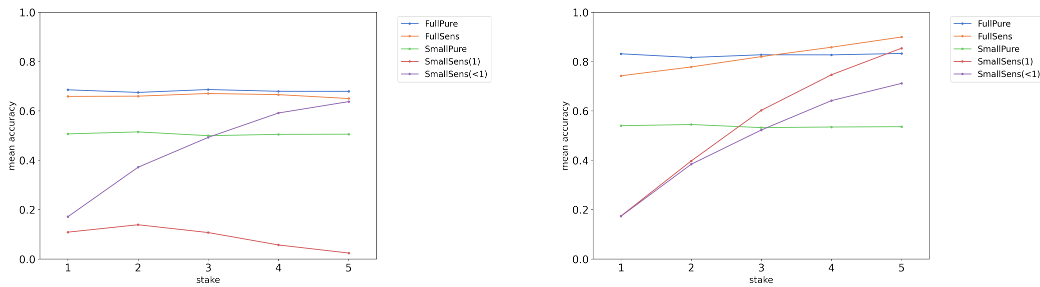
Taking stock, I have categorized the family of threshold views into four types, characterized by two orthogonal features pertinent to the action-guiding role of outright beliefs. In the class of time-sensitive decision problems modelled here, no single type of threshold view performs well across different contexts. A good cognitive strategy cannot squarely fit into any of the four types. In table 2, I summarize the strengths and weaknesses of each type of threshold view.

5 Discussions

5.1 Cognitive Flexibility

The results of the simulation shed light on which cognitive strategies lead to good action-guiding beliefs in decision contexts characterized by different stakes and difficulty. The upshot is that flexibility is an important asset. The question of which cognitive strategy is the best cannot be answered in isolation to the context of deliberation. The best-performing agent is

[25]The results do not render FullSens, specifically in high stake situations, a failed strategy *tout court*, since there might be situations where errors are so severely penalized that FullSens agents' level of prudence in high-stake scenarios is warranted. This is not the case in the present setup.



(a) All Trials

(b) Only trials with outright beliefs formed

Figure 5: (a) The correlation between stakes and accuracy in all trials. (b) The correlation between stakes and accuracy in trials where an outright belief is formed within 100 time steps.

	Epistemic Purist	Pragmatic Sensitivist
Full Domain	• Performs well in easy contexts, in particular high-stake ones. • Performs poorly when the problem is difficult.	• Performs well in easy contexts, in particular low-stake ones. • Performs poorly in difficult, high-stake context.
Small Domain	• Performs especially well in difficult high-stake contexts.	• Performs well in difficult low-stake contexts, as well as easy high-stake ones. • Performs poorly in easy, low-stake contexts.

Table 2: A summary of the strengths and weaknesses of different types of threshold views in time-sensitive decision problems.

not of a specific type, but instead one that is able to flexibly adjust their cognitive strategies based on different contexts. Even with all the simplifications made in the simulations, no single strategy can dominate in all contexts. This makes a strong case that cognitive flexibility is crucial in real life situations with far higher complexity.

Given the flexibility required for an agent to perform well in problems with time constraints, we should reconsider whether the threshold view really provides a simple algorithm that streamlines our reasoning and decision making. On the one hand, if the threshold view can provide a path from credences to binary beliefs simple enough to fit into one of the four categories, the resulting outright beliefs will be unlikely to guide our actions well in a diverse range of decision problems. On the other hand, if we were to have a threshold view that allows for enough flexibility to adapt to different decision contexts, then it is not clear how much we would gain in terms of streamlining our reasoning, calling the advertised simplicity of the threshold view into question.

Like in all models, some arbitrary choices had to be made in the simulation (e.g., the standard threshold of 0.9). If we change the arbitrarily set parameters enough, then there is a chance that the outcomes and relative performances of different types of agent might not remain constant. However, I doubt that this will affect the upshot discussed here. My arguments are based on variation and not the exact order of relative performances in the results. All that is needed for the worry to remain is for different threshold strategies to be suitable for different decision contexts. I suspect that if we change the setup so much that there is no longer variation, then the simulation will no longer model anything remotely realistic.

Staffel (2019) highlights the tension between the supposed function of outright beliefs to simplify reasoning and the cognitive mechanisms required to switch between graded and binary beliefs. She argues that the computation required to maintain coherence when moving between credences and outright beliefs undermines the purpose of such a transition to save on computation. In a sense, my worry echoes hers, though my focus here is on the action-guiding function of outright beliefs, while hers is on ideal norms of coherence governing beliefs.

5.2 Adaptation of Search Termination

Research in cognitive psychology have long embraced cognitive flexibility, as is evidenced by important work on the adaptation of search termination.

> Inherent in having a capability to search is having a capability to terminate search, and an ability to learn, adapt and regulate the termination of search. A search process that does not terminate is not useful, and termination strategies that are unable to learn from experience, or adapt to changes in the environment or the goals of a decision maker will usually be inefficient and limiting. (Lee et al., 2014, p. 245)

Given that forming an outright belief may be seen as a sort of search termination, research in cognitive psychology can inform us about how our strategies to form outright beliefs adapt to different decision contexts. The empirical work by Lee et al. (2014), for example, examines how human agents change their search termination strategies in response to changes in the environment, and Newell (2005) investigates how we decide which cognitive strategy to employ under varying environmental constraints. While computer simulations facilitate our understanding of the normative dimension of search termination, psychological experiments offer direct evidence for descriptive theories. Models in cognitive psychology can also inform us on how to build more realistic models in the simulations.

5.3 Future Directions

The scope of this study is limited. I focused on formulations of the threshold view that have been proposed in the literature, plus a few other variations (which, interestingly, outperformed the models proposed in the literature in many cases) for the completeness of the conceptual space. Other variations of the threshold view, such as alternative specifications of SmallSens, different ways the family of threshold views can be categorized and other decision problems remain to be studied. For instance, we might add some costs to cognitive resources such as memory storage.

Another direction to explore is how stakes can be defined differently. In the current simulations, stakes are added as an additional parameter, but we can also construct stakes based on the existing parameters that define how the probability of death changes with time for patients. The shape and scale of the cumulative density function of Weibull distributions influence how urgent it is for the doctors to make a decision. When the curve is steeper, the probability of death increases faster, so the doctor is under more pressure to act quickly. If the correct intervention drastically reduces the probability of death and the wrong intervention increases the probability of death by a large margin, the sensitivist agent may in response be more prudent. Generating the stake of each case based on these factors may yield new insights.[26] Small-domain agents might also adjust their domain sizes partially based on how concentrated their prior probabilities are to reduce the number of active hypotheses without greatly increasing the chance that the truth is excluded from the domain.

A Weibull Distribution

The Weibull distribution is a continuous probability distribution defined by the function

$$f(x;\lambda,k) = \begin{cases} \frac{k}{\lambda}(\frac{x}{\lambda})^{k-1}e^{-(x/\lambda)^k}, & \text{if } x \leq 0. \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

where $\lambda \in (0, +\infty)$ is the scale parameter and $k \in (0, +\infty)$ is the shape parameter. The cumulative distribution function of the Weibull distribution is given by

$$f(x;\lambda,k) = \begin{cases} 1 - e^{-(x/\lambda)^k}, & \text{if } x \leq 0. \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

B Symmetric Dirichlet Distribution

The Dirichlet distribution is a multivariate generalization of the beta distribution, parameterized by a vector $\boldsymbol{\alpha}$ with real entries. The Dirichlet distribution with parameters $\alpha_1, \alpha_2, ..., \alpha_K$ is given by

$$f(x_1, x_2, ..., x_K; \alpha_1, \alpha_2, ..., \alpha_K) = \frac{1}{B(\boldsymbol{\alpha})} \prod_{i=1}^{K} x_i^{\alpha_i - 1}, \tag{4}$$

[26]This suggestion requires the further assumption that an agent knows the costs of type 1 and type 2 errors, as well as how the costs change with time.

where $\sum_{i=1}^{K} x_i = 1$ and $x_1, x_2, ..., x_K > 0$. The normalizing constant $B(\boldsymbol{\alpha})$ is the multivariate beta function

$$B(\boldsymbol{\alpha}) = \frac{\prod_{i=1}^{K} \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^{K} \alpha_i)}, \tag{5}$$

where Γ is the gamma function.

A special case of the Dirichlet distribution is the *symmetric* Dirichlet Distribution. In a symmetric Dirichlet distribution, the entries of the vector $\boldsymbol{\alpha}$ are all equal. This corresponds to the case where there is no prior information that leads us to prefer one possibility over another. Since all the entries of $\boldsymbol{\alpha}$ are the same, the symmetric Dirichlet distribution can simply be parameterized by a scalar value α, often called the concentration parameter. When $\alpha > 1$ the probability is randomly but evenly distributed across the possibilities. When $\alpha < 1$, most of the probability will be concentrated in a few of the possibilities. The symmetric Dirichlet distribution is given by

$$f(x_1, x_2, ..., x_K; \alpha) = \frac{\Gamma(\alpha K)}{\Gamma(\alpha)^K} \prod_{i=1}^{K} x_i^{\alpha_i - 1}. \tag{6}$$

C Code for the Model

The code for the model, written in Python using the agent-based simulation package Mesa, can be found here: https://github.com/alicecwhuang/threshold_code.git.

Acknowledgement

I am grateful to Jonathan Weisberg, Jennifer Nagel, Boris Babic as well as two anonymous referees for their helpful comments on various drafts of this paper. Special thanks are also due to Titas Geryba for much-appreciated copy-editing and insightful suggestions. Thanks also to Nischal Mainali for interesting interdisciplinary discussions about cognitive constraints.

References

B. Babic. A theory of epistemic risk. *Philosophy of Science*, 86(3):522–550, 2019. doi: 10.1086/703552.

D. Christensen. Clever bookies and coherent beliefs. *Philosophical Review*, 100(2):229–247, 1991. doi: 10.2307/2185301.

R. Clarke. Belief is credence one (in context). *Philosophers' Imprint*, 13:1–18, 2013.

N. Cowan. What are the differences between long-term, short-term, and working memory? In W. S. Sossin, J.-C. Lacaille, V. F. Castellucci, and S. Belleville, editors, *Essence of Memory*, volume 169 of *Progress in Brain Research*, pages 323–338. Elsevier, 2008. doi: https://doi.org/10.1016/S0079-6123(07)00020-9.

N. Cowan. The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, 19(1):51–57, 2010.

N. Cowan, C. C. Morey, Z. Chen, A. L. Gilchrist, and J. S. Saults. Theory and measurement of working memory capacity limits. In B. H. Ross, editor, *Advances in Research and Theory*, volume 49 of *Psychology of Learning and Motivation*, pages 49–104. Academic Press, 2008.

I. Douven. The ecological rationality of explanatory reasoning. *Studies in History and Philosophy of Science Part A*, 79:1–14, 2020. doi: 10.1016/j.shpsa.2019.06.004.

R. A. Duckworth, A. L. Potticary, and A. V. Badyaev. On the origins of adaptive behavioral complexity: Developmental channeling of structural trade-offs. In M. Naguib, L. Barrett, S. D. Healy, J. Podos, L. W. Simmons, and M. Zuk, editors, *Advances in the Study of Behavior*, volume 50 of *Advances in the Study of Behavior*, pages 1–36. Academic Press, 2018.

K. Easwaran. Dr. truthlove or: How i learned to stop worrying and love bayesian probabilities*. *Noûs*, 50(4):816–853, 2016.

J. Fantl and M. McGrath. Evidence, pragmatics, and justification. *Philosophical Review*, 111(1):67–94, 2002. doi: 10.1215/00318108-111-1-67.

R. Foley. *Working Without a Net: A Study of Egocentric Epistemology.* New York: Oxford University Press, 1992.

K. Frankish. *Partial Belief and Flat-Out Belief*, pages 75–93. Springer Netherlands, Dordrecht, 2009.

J. Friedman. Inquiry and belief. *Noûs*, 53(2):296–315, 2019. doi: 10.1111/nous.12222.

D. Ganson. Evidentialism and pragmatic constraints on outright belief. *Philosophical Studies*, 139(3):441–458, 2008. doi: 10.1007/s11098-007-9133-9.

J. Gao. Credal pragmatism. *Philosophical Studies*, 176(6):1595–1617, 2019. doi: 10.1007/s11098-018-1081-z.

J. Gao. Credal sensitivism: Threshold vs. credence-one. *Inquiry: An Interdisciplinary Journal of Philosophy*, forthcoming.

G. A. George A. Miller, E. Galanter, and K. H. Pribram. *Plans and the structure of behavior.* Holt, Rinehart and Winston, 1960.

D. Greco. 4. probability and prodigality. *Oxford Studies in Epistemology*, 4:82, 2013.

D. Greco. How i learned to stop worrying and love probability 1. *Philosophical Perspectives*, 29(1):179–201, 2015. doi: 10.1111/phpe.12059.

J. Hawthorne. *Knowledge and Lotteries.* Oxford University Press, 2003.

J. Hawthorne and J. Stanley. Knowledge and action. *Journal of Philosophy*, 105(10):571–590, 2008. doi: 10.5840/jphil20081051022.

R. Holton. Intention as a model for belief. In M. Vargas and G. Yaffe, editors, *Rational and Social Agency: Essays on the Philosophy of Michael Bratman.* Oxford University Press, 2014.

H. E. Kyburg. *Probability and the Logic of Rational Belief*. Wesleyan University Press, 1961.

M. D. Lee, B. Newell, and J. Vandekerckhove. Modeling the adaptation of search termination in human decision making. *Decision*, 1(4):223–251, 2014.

P. Maher. *The concept of acceptance*, page 130–161. Cambridge Studies in Probability, Induction and Decision Theory. Cambridge University Press, 1993. doi: 10.1017/CBO9780511527326.007.

O. Mayseless and A. W. Kruglanski. What makes you so sure? effects of epistemic motivations on judgmental confidence. *Organizational Behavior and Human Decision Processes*, 39(2):162–183, 1987. doi: https://doi.org/10.1016/0749-5978(87)90036-7.

J. Nagel. Epistemic anxiety and adaptive invariantism. *Philosophical Perspectives*, 24(1):407–435, 2010. doi: 10.1111/j.1520-8583.2010.00198.x.

B. R. Newell. Re-visions of rationality? *Trends in Cognitive Sciences*, 9(1):11–15, 2005. doi: 10.1016/j.tics.2004.11.005.

A. Norby. Uncertainty without all the doubt. *Mind and Language*, 30(1):70–94, 2015. doi: 10.1111/mila.12072.

J. Ross and M. Schroeder. Belief, credence, and pragmatic encroachment. *Philosophy and Phenomenological Research*, 88(2):259–288, 2014. doi: 10.1111/j.1933-1592.2011.00552.x.

K. Rubin. Total pragmatic encroachment and epistemic permissiveness. *Pacific Philosophical Quarterly*, 96(1):12–38, 2015. doi: 10.1111/papq.12060.

J. Staffel. How do beliefs simplify reasoning? *Noûs*, 53(4):937–962, 2019. doi: 10.1111/nous.12254.

J. Stanley. *Knowledge and Practical Interests*. Oxford University Press, 2005.

J. Stanley. Replies to gilbert harman, ram neta, and stephen schiffer. *Philosophy and Phenomenological Research*, 75(1):196–210, 2007. doi: 10.1111/j.1933-1592.2007.00071.x.

B. Weatherson. Can we do without pragmatic encroachment. *Philosophical Perspectives*, 19(1):417–443, 2005. doi: 10.1111/j.1520-8583.2005.00068.x.

J. Weisberg. Belief in psyontology. *Philosophers' Imprint*, 20(11), 2020.

T. Williamson. *Anti-Luminosity*, page 93–114. Oxford University Press, 2000.