

Betting on Conspiracy:

A Decision Theoretic Account of the Rationality of Conspiracy Theory Belief

Melina Tsapos

Abstract

The question of the rationality of conspiratorial belief divides philosophers into mainly two camps. The particularists believe that each conspiracy theory ought to be examined on its own merits. The generalist, by contrast, argues that there is something inherently suspect about conspiracy theories that makes belief in them irrational. Recent empirical findings indicate that conspiratorial thinking is commonplace among ordinary people, which has naturally shifted attention to the particularists. Yet, even the particularist must agree that not all conspiracy belief is rational, in which case she must explain what separates rational from non-rational conspiratorial thinking. In this paper, I contrast three strategies to this end: 1) the probabilistic objectivist, who assesses the objective probability of conspiracies; 2) the subjectivist, who rather focuses on the perspective of the believer, and typically views the decision to believe in a conspiracy as a problem of decision making under risk. Approaches 1) and 2) rely on assessments of the probability of conspiracy which, I argue, limits their applicability. Instead, I explore 3) viewing the problem facing the potential believer as a decision problem under uncertainty about probabilities. I argue, furthermore, that focusing solely on epistemic utilities fails to do justice to the particular character of conspiracy beliefs, which are not exclusively epistemically motivated, and I investigate the rationality of such beliefs under a number of standard decision rules.

1. Introduction

The rationality of belief in conspiracy theories is a hotly debated issue in contemporary philosophy. There are two main views in the debate. There are the generalists who argue that belief in conspiracy theories is *prima facie* irrational (Cassam, 2019; Napolitano, 2022; Napolitano and Reuter 2021; Sunstein and Vermule, 2009). By contrast, the particularists understand conspiracy theories to be putative explanations and argue that each conspiracy theory ought to be examined on its own merits. According to the particularist, we cannot assess conspiracy theories as a class but, rather, conspiracy theories should be assessed on a case-by-case basis (Basham, 2018; Dentith,

2016; 2019; Coady, 2012; Keeley, 1999). On this view, there is nothing inherently irrational about belief in conspiracy theories; after all, conspiracies unravel every day, and most person would agree that history is full of secretive plots, political and otherwise (Pigden, 1995).

However, empirical research indicates that “conspiracy beliefs emerge as ordinary people make judgments about the social and political world” due to situational triggers and subtle contextual variables (Radnitz and Underwood, 2015, p. 113). Reviewing recent scientific findings, Brotherton (2015) concludes, in the same vein, that conspiracy belief is far from being generally bizarre, but that “[w]e have innately suspicious minds” and we are all “natural-born conspiracy theorists” (ibid., p.17). Poth and Dolega (2023) show that conspiracy belief can be Bayes-rational. It is true that the fact that people are naturally-born conspiracy theorists is logically compatible with generalism about conspiracy theories. However, there is a tension or incoherence between these ideas. How come people are so prone to conspiracy belief if such belief is *prima facie* suspect? The tension is less obvious between people being naturally-born conspiracy theorists and particularism about conspiracy theories. If some conspiracy theories are plausible, it comes as less of a surprise that people are inclined to believe in them. However, this raises the problem of explaining what separates rational from non-rational belief in conspiracy theories for the particularist.

In this paper I will contrast three strategies to this end. First, I will present the objectivist strategy, which assesses the objective probability of conspiracy theories. Second, I consider the subjectivist approach, which focuses on the rationality of conspiracy belief from the point of view of the believer. There are two standard perspectives: under risk and under uncertainty. First, I will consider the decision to believe in a conspiracy in the framework of decision making under risk using determinate probabilities and epistemic utilities. However, relying on determinate probabilities seriously limits the applicability of both these approaches. Moreover, focusing solely on epistemic utilities fails to do justice to the particular character of conspiracy beliefs, which are not exclusively epistemically motivated. Thus, lastly, I explore the second perspective of the

subjectivist approach, viewing belief in conspiracy theories under uncertainty about probabilities. In this context, I investigate the rationality of conspiracy belief under a number of standard decision rules and sources of non-epistemic utilities.

2. The objectivist

The objectivist view holds that external factors of a conspiracy theory determine its rationality. Thus, the objectivist places his or her reliance on empirical data to determine the viability of conspiracy theories. For instance, a large conspiracy may be more likely to be revealed as time passes, and the corresponding theory becomes increasingly unlikely when the conspiracy is not revealed. This is precisely what David Robert Grimes argues (Grimes, 2016: 2). Following his definition, a conspiracy theory is an explanation of some event or practice that references the machinations of a (usually powerful in some sense) small group of people, who attempt to conceal their role (at least until their aims are accomplished).¹ Grimes acknowledges that there are conspiracy theories that are viable and therefore believing in them would not be irrational, writing that it would be “unfair [to] dismiss all allegation of conspiracy as paranoid where in some instances it is demonstrably not so”. Grimes goes on to suggest that there “must be a clear rationale for clarifying the outlandish from the reasonable”. His suggestion is that we can determine the likelihood of the viability of a conspiracy theory by looking at the number of conspirators involved, as well as the amount of time passed since the conspiracy occurred.

To this end, Grimes puts forth a simple mathematical model for estimating just such viability. The parameter p is the probability of an intrinsic leak or failure, and is estimated from

¹ To be exact, Grimes defines a conspiracy theory as “an effort to explain some event or practice by reference to the machinations of powerful people, who attempt to conceal their role (at least until their aims are accomplished)” (2016: 1-2). Since a theory is hardly an “effort”, I have amended this part of his definition above. For other definitions and for a further discussion on the importance of the definition, see Tsapos, (2023); Napolitano and Reuter (2021).

known conspiracies, what Grimes calls “exposed examples” (6). The three examples he uses are the National Security Agency PRISM affair, secretly conducting excessive surveillance of internet traffic; the Tuskegee syphilis experiment, where the US Public Health Services made unethical observational studies on African-American men; and the FBI forensic scandal, the pseudoscientific nature of the bureau’s forensic testing resulting in many innocent men being incarcerated. Taking the number N of conspirators and the life-time t of the conspiracy and calculating the failure odds, Grimes suggests that we can predict “a best-case scenario” for a conspiracy theory’s viability. In his model, Grimes assumes that everyone involved is in on it. For example, in the NSA case, Grimes assumes that all employees at NSA at the time, approximately 30,000, were conspiring. He makes a number of further assumptions: 1) that the conspirators are dedicated to keeping their conspiracy a secret, 2) that leaks by the conspirators expose the conspiracy, 3) that the number of conspirators over time, will depend on one of the following factors: If the conspiracy requires constant upkeep, then the number of conspirators remains constant; if the conspiracy is a single event and no new conspirators are required, then the number of conspirators involved will decay over time; and, finally, if the conspirators are rapidly removed for some reason, there is an exponential decay. I will return to these matters in a moment.

With data from these three cases above, Grimes estimates the probability of failure due to internal leaks, with the lower bound given by²

$$p > 1 - \sqrt[N(t)]{1 - \frac{\ln 2}{t}}$$

² Grimes acknowledges the considerable and “unavoidable ambiguities” that his estimates rely on. For example, the total number of staff compared to the number of people with full knowledge of the event; and complicating factors such as estimating the life-time of a conspiracy, which is not always clear. For example, in the Tuskegee experiment case, the original experiment commenced in the 1930s, but didn’t become unethical until the late 1940s. I will return to some concerns in the text. For all the ambiguity concerns raised by Grimes, see Grimes (2016).

Thus, the probability of failure is seen as a function of N and t , such that a high probability yields a low chance of a successful conspiracy.

With Grimes' model in mind, consider the following example. The Swedish prime minister Olof Palme was assassinated in central Stockholm on February 28, 1986, while walking home one night from the movies. The assassination and subsequent police investigation have caused one of the greatest controversies in Swedish history. The case was considered closed by the authorities after more than 35 years of investigations, a decision that faced heavy criticism. Many explanations as to what really happened are still floating around, among them conspiracy theories per the Grimesian definition. For example, a version of the so-called SÄPO-conspiracy³ identifies the Swedish Security Service (SÄPO), together with its American counterpart, CIA, as colluding to assassinate the prime minister. According to Grimes' analysis, we should be able to predict the intrinsic failure of this theory by estimating the number of actors involved after a certain time, in this case about 1,400 working for SÄPO and roughly 21,000 for the CIA⁴. Using the above equation, the estimate time for such conspiracy to be revealed would be approximately 34 years⁵.

³ Högström, (2014).

⁴ I have used estimates from 2022 (and not from 1986) for the sake of using readily accessible, and accurate data, and to better illustrate Grimes' equation. I have also assumed that, on the SÄPO-conspiracy, the time passed from conspiracy to murder was negligible.

⁵ The estimated failure time is based on Grimes' Vaccination Conspiracy example (which shares similar estimation of the number of conspirators involved to the SÄPO-conspiracy example) to avoid having to further speculate about factors such as the several possibilities for the parameter $N(t)$, the age of the conspirators and other factors necessary to estimate the p -value for this case specifically. In any case, the SÄPO-conspiracy example is only meant to illustrate Grimes' model, and any data used here has not been extensively verified (in any case, much of the necessary data to establish the p -value is unfortunately not public or accessible).

Since more than 36 years have passed since the murder took place, we can rule out the SÄPO-conspiracy as a non-viable conspiracy theory.

The objectivist view, then, seems to offer a clear-cut way of determining what is a viable conspiracy theory. All we need is to plug in some figures and we can end the discussion once and for all. However, even though this approach is indeed important in many ways, and Grimes may very well be correct in proposing that the number of conspirators has an effect on it potentially being leaked to the public, it suffers from a number of challenges that are not so easily overlooked.

One criticism raised by Dentith (2019: 17) targets the examples that Grimes uses for the parameter estimation, more specifically the manner in which the three conspiracies were exposed. Grimes has it that a leak generally comes from a conspirator either accidentally (where a conspirator fails in some way to cover up the conspiracy) or intentionally (where a whistleblower purposefully leaks the information). However, Dentith argues that the whistleblower Edward Snowden, for instance, in the NSA case was an outsider, i.e., someone who discovered the conspiracy without having been part of it. Similar observations apply to the two other exposed conspiracies. Dentith shows that in all three examples, the leak originated from outsiders. Thus, Dentith concludes that there is “a mismatch between Grime’s chosen examples, and his theory about how leaks over time revealed and made these conspiracies redundant; his examples fail to capture the very thing he wants to measure”. Dentith further criticize Grimes’ account on the grounds of its failure to distinguish between conspirators and whistleblowers, stating that it is even possible to be involved in a conspiracy without realizing you are conspiring, in which case you would not see the need to become a whistleblower.⁶

⁶ Dentith further elaborates this point, pointing out that some people are dupes or patsies being used by the core set of conspirators. They may not know that they are involved in a conspiracy because they do not need to know the plot. As such, they can end up revealing the conspiracy just by talking, because they don’t know they are in it and need to keep things secret. They are also not necessarily interested in exposing it. Either way, Dentith’s point is a valid one.

While I agree with Dentith's criticism, in the present context I believe there is another fundamental concern, one that I take Grimes himself anticipates. He writes: "There is also an open question of whether using exposed conspiracies to estimate parameters might itself introduce bias and produce overly high estimates of p – this may be the case, but given the highly conservative estimates employed for other parameters, it is more likely that p for most conspiracies will be much higher than our estimate, as even relatively small conspiracies (such as Watergate, for example) have historically been rapidly exposed" (9). However, Grimes severely underestimates the significance of this problem. I will briefly elaborate this point.

The working definition of conspiracy theories does not itself *a priori* dismiss conspiracy theories as inherently false. Let us call the class of all conspiracy theories CT. Even so, there are conspiracy theories that we may dismiss, because, according to Grimes, they are "demonstrably nonsensical" and outlandish; I call this class O. Typically, they contradict scientific thinking or are based on evidence incompatible with science. Then there are those conspiracy theories that, according to Grimes are reasonable and whose viability is still an open question; call this class R. These conspiracy theories "would be unfair to simply dismiss", because they reflect real concerns of possible collusion. Importantly, they build on a narrative that can be challenged by scientific evidence. We can then estimate using Grimes' model which of these are viable after time t ; I shall call this class V. The class of conspiracy theories that are considered reasonable, but are not viable after time t , I will call NV. Yet, another class of conspiracies already mentioned are exposed conspiracies, E.

According to Grimes, the model provides an estimate for the intrinsic failure of reasonable conspiracies after some time t . Importantly, these conspiracies may or may not be exposed in the future. However, the parameter estimates in the model are based solely on exposed conspiracies.

Mixing these duped people involved in the conspiracy that may or may not unknowingly reveal the conspiracy with the conspirators seems rather counterintuitive.

Now, this matters because, as Grimes (2016: 3-4) himself indeed emphasizes, “conspirators are in general dedicated for the most part to the concealment of their activity” so that “the act of a conspiracy being exposed is a relatively rare and independent event”. In fact, one of Grimes’ assumptions c), about a decaying group of conspirators, lend further support to the latter claim. In other words, Grimes’ parameter estimates introduce not only “overly high estimates” of p , but quite possibly intolerably so. Another way of putting it, is that the model does not estimate the probability of failure *per se*, but the conditional probability of failure given that the conspiracy will be exposed at some point.

In order to make probability estimates about the intrinsic failure of conspiracy theories generally, a model must be based on a representative sample from R , one that is not skewed toward conspiracies that have been or will be exposed. Granted, this is very difficult, and most likely an unrealistic enterprise. In effect, this is a dilemma for Grimes: either the model quite possibly grossly overestimates the probability of intrinsic failure or it provides no guidance whatsoever as to the viability of a given reasonable conspiracy theory.

Nevertheless, important lessons are still to be had from the objectivist account, for example concerning the impact of the size factor when estimating the likelihood of conspiracy theories. But just *how* size matters is still an open question. Since Grimes’ model in the end cannot tell us much about the rationality (or lack thereof) of believing a conspiracy theory, I now turn to an alternative, subjectivist account.

3. The subjectivist: believing in conspiracy as a decision under risk

The subjectivist framework looks at belief in conspiracy theories as they arise from rational Bayesian cognition and considers the individual believer’s perspective, assuming that a decision-maker seeks to maximize the payoffs of their decision. I will use Gabriel Doyle (2021)’s account for detailing this framework. The account builds on previous research by Olsson (2020) indicating

that seemingly irrational belief polarization can arise from ideally rational Bayesian reasoning, where two ideal Bayesian agents can end up divided on an issue given the exact same evidence to consider. Doyle believes that something similar is true for belief in a conspiracy theory: such seemingly irrational belief can arise from impeccably rational processes of belief formation.⁷

In the spirit of Bayesianism, Doyle assumes that a person seeks the most satisfying explanation H for a set of observed phenomena D based on the posterior probabilities of the relevant hypotheses. These probabilities are calculated as the product of the likelihoods and priors for each hypothesis. To determine which hypothesis is preferred, Doyle looks at the epistemic utility of choosing a high-posterior hypothesis together with other additional utilities derived from the hypothesis themselves. In his framework there are two primary explanations available: the conspiracy hypothesis, H_{CT} , and the non-conspiracy theory hypothesis, H_{NC} . The posterior probability for each hypothesis is proportional to the likelihood of seeing observed data under a given hypothesis, times the prior belief in that source. The utility is in Doyle's (2021: 3) framework "primarily derived by identifying the true source of the data". This "epistemic" utility is a function of the posterior probability of the hypothesis given the data:

$$U_{epistem}(H;D)=\log p(H|D)\propto\log p(D|H)+\log p(H)$$

Doyle considers that there may be other utilities factoring into the decision-making, such as personal satisfaction from having a belief. I will return to this topic in the next section. Further, he elaborates on a number of ways in which normatively optimal reasoning could lead someone to support or commit to a conspiracy theory, such as having elevated prior beliefs in conspiracies;

⁷ In his book *Bad Beliefs* (2021) Neil Levy makes a similar claim to Doyle, and argues that ordinary people are rational agents whose beliefs are the result of their rational response to evidence they're presented with. However, according to Levy there is higher-order evidence, such as trust, that are typically neglected but should be recognized and managed in order to tackle bad beliefs.

having different likelihood functions, and having missing hypotheses. I will summarize each of these in turn.

According to the elevated prior hypothesis, strong beliefs in conspiracies raise the individual’s prior belief in a conspiracy theory explanation for future phenomena that may present themselves. According to Doyle, the hypothesis coincides with other findings about conspiracy beliefs, for example that conspiracy theory beliefs “appear to cluster”, (which I take to mean that people tend to believe more than just one conspiracy theory that are structurally similar)⁸. If two people agree on a given CT’s likelihood and non-epistemic utility, the elevated conspiracy theory prior may be enough to yield a higher posterior for a conspiracy theory believer. Table 1 from Doyle (2021: 4) indicates that with a change from 5% to 20% in prior expectation of conspiracy, a rational non-conspiracy theory believer (N) is converted into a conspiracy theory believer (B).

	H	$p(H)$	$p(D H)$	$U(H)$	$U(H D)$
B	H_{CT}	0.2	0.1	0	-3.91
	H_{NC}	0.8	0.9	-4	-4.33
N	H_{CT}	0.05	0.1	0	-5.30
	H_{NC}	0.95	0.9	-4	-4.16

Table 1: Doyle’s table of elevated priors. B is CT- believer’s probability, N is non-believer’s. H_{CT} is the CT, H_{NC} is the non-CT explanation.

Doyle next consideration is the different likelihood function, which state that even if two people agree on the set of hypotheses, the observed data and the conditional probability distribution, they may still end up with different posteriors, if they disagree on the value of another factor, for

⁸ Swami et al., (2011) and Wood et al., (2012) are two examples of research studies that support this idea that conspiracy theories operate within “monological belief systems”, in which conspiracy theorists find support for conspiratorial beliefs in other conspiratorial beliefs, or in related generalizations. However, see Hagen 2018 that argues that this is either wrong or at least misleading.

example which experts they trust. Take climate-change denial as an example. A person's likelihood function for the scientific consensus on climate change depends according to Doyle on a third variable, namely the trust placed on scientists. This factor may be enough to change the posteriors. In such cases, Doyle argues, the probability of the conspiracy theory hypothesis is higher for the conspiracy theory believer, who thinks scientist would lie to maintain a desired consensus than for the non-believer who considers this possibility unlikely. The third explanation of conspiracy belief is the missing hypothesis explanation. In this case the source of accepting conspiracy theory beliefs is that the true non-conspiracy explanation is not part of the conspiracy theory believer's arsenal of potential explanations.

While Doyle should be credited for the subjectivist approach to conspiracy belief and for noticing, without further elaboration, the potential role of non-epistemic utilities, I believe this framework needs to be further developed, in order to better understand conspiracy beliefs. My reasons for this are threefold: First, more can be done about the fact that a decision is made under conditions of *known probabilities*, which makes a framework such as Doyle's vulnerable to the same problems as the objectivist account, namely that such known probabilities are difficult to come by (see previous section), which is of course also a worry for decision theory generally; second, it relies on numerical utilities, which again, calls for speculation as to what to base those numbers on; and third, I argue, it doesn't capture the non-epistemic utilities involved in many conspiracy beliefs. These remarks suggest that an alternative view is worth developing in addition to the already existing frameworks, namely accounting for conspiracy belief within decision theory under uncertainty about probabilities, focusing on non-epistemic utilities, which I will develop next.

4. The subjectivist: believing in conspiracy as a decision under uncertainty

Belief in conspiracy theories⁹ can sometimes be primarily psychologically motivated at the expense of a primary concern for truth. As such, they can fill the need to explain unusual or stressful events such as terrorist attacks, disease outbreaks, scandals, illegal mass surveillance and sudden, complex deaths (van Prooijen and van Vugt, 2018). They may reduce the complexity of uncertainty into focused fear, and they may also motivate collective action, when used in the context of relationships between in and out groups (Franks, 2013). Conspiracy theories can serve multiple purposes, including various political grievances (including legitimate ones) in contemporary society (Doyle, 2021; Gosa, 2011). As such, the conspiracy theorist emerges as a social phenomenon, and not a cognitively challenged or epistemically vicious outlier. This should arguably be reflected in a decision theory about our ordinary choice situations. Typically, we are not handed probabilities about the possible states the world might be in. Rather, as decision makers we often have to make our decisions under uncertainty. Considering the notion of non-epistemic utility coupled with the fact that we do not often have posterior probabilities for our choices, can provide a broader framework for understanding conspiracy theories, and why rational people believe in them.

I will consider two potential sources of non-epistemic value, which I call *puzzle hunt* and *social power*. Puzzle hunt has two aspects: one is the human instinct to solve puzzles, and the social effects that follows. For example, one might gain and increased social status, social credit or other form of recognition, by solving a mystery or puzzle. Social power involves ways in which a conspiracy theory belief may contribute to a social power struggle. It may manifest itself in protesting against oppressors, or undermining authority by challenging the reputation

⁹ What I say in the following is equally true of other kinds of attitudes and endorsements of conspiracy theories. It does not specifically rely on the attitude of believing. For an extensive discussion on non-doxastic attitudes towards conspiracy theories, see Ichino and R aikka, (2021). See also section 5.

of those believed to have conspired.¹⁰ I will discuss each in turn and I will apply a number of decision rules to formalize the circumstances under which a rational decision-maker may come to accept a conspiracy theory over a non-conspiracy theory. Of course, these are only two sources of non-epistemic value, and not meant to be an exhaustive list.

4.1 Puzzle hunt

Humans like a good puzzle, and as the anthropologist Marcel Danesi argues, the need to solve puzzles, for mystery and suspense, is displayed in humans at a level not found in any other animal species. Danesi (2002: 2) writes that “puzzles and mysteries are intrinsically intertwined in human life” since the dawn of history; a phenomenon I shall dub *Puzzle hunt*. The question of why people are fascinated by seemingly trivial puzzles that often require substantial time and mental effort to solve is beyond the scope of my discussion here. However, I would suggest that there is a social aspect to puzzle hunt, much like the concept of “counterknowledge” and “treasure hunt” described by Gosa, (2011). Puzzle hunt is social in that it can be socially motivated, by the desire to gain appraisal for being the one who solved a mystery.

Puzzlehunt, then, as characterized above, is arguably driving some conspiracy theorist. Considering again the case of the conspiracy theories surrounding the assassination of the Swedish prime minister. Many felt dissatisfied when years of police investigations, only led to the wrongful arrest and conviction of a suspect who was later freed. In light of these events, a

¹⁰ The flip-side is of course that the oppressor utilizes conspiracy theories in order to keep the social power. Either way I believe it is two sides of the same coin. Both the oppressor and the oppressed must view the conspirator as powerful enough to be a threat, and the motivations for committing to a conspiracy theory would be similar in nature (such as undermining the alleged conspirators).

group was formed that became known as “the private investigators” (or “privatspanarna”). Some of the private investigators organized regular meetings, discussing and debating the puzzle-pieces they felt were missing to solve the murder-mystery. Some of the private investigators were retired police officers and journalists that, even after retiring, continued working on the case. What they all had in common was a feeling that the mystery was not yet solved and that they wanted to find out what *really* happened that winter night in 1986. The concluding remarks on the website palmemordet.se leaves us with the same sentiment: “[the] murder of Olof Palme remains a mystery with no sure answers”¹¹.

Let us now consider the problem of deciding which hypothesis to believe from the perspective of puzzle hunting. To fix ideas, we consider the conspiracy theory (CT) to be about the Palme murder, and the decision-maker might be a member of the group of private investigators. There are two possible states of the world: either CT is true or CT is false. The decision maker can either believe in CT (B_{CT}) or not believe in CT (B_{NC}). Given a possible state of the world, the outcome of believing the one or the other, in terms of how the decision-maker is likely to be perceived by her peers, will be different, as illustrated in Table 2.

Puzzle hunt	<i>CT true</i>	<i>CT false</i>
B_{CT}	U1: “First to solve the puzzle”	U2: “A crazy conspiracy theorists”
B_{NC}	U3: “Naïve, duped by conspirators”	U4: “Not misled. Being reasonable”

Table 2. Potential utility preferences of a private investigator, with *Puzzle hunt* as a source for non-epistemic value.

¹¹ In the original: “Mordet på Olof Palme förblir en gåta utan säkra svar”. <https://www.palmemordet.se/> Accessed on 2 dec 2022.

According to table 1, U1 is the most prioritized utility due to the great satisfaction of pursuing a puzzle hunt and being the one who was able to solve it, thus being right. However, if one believes the conspiracy theory and it is false, there is a greater negative loss, compared to not believing the conspiracy given the same outcome. There are a number of reasons for this, one of which can be a loss of social status when a theory is believed and otherwise acted upon (by telling others or in other ways expressed through behavior), and there is further risk that people will not believe or take the persons conviction seriously in the future. In the scenario where one does not believe the conspiracy theory and it is true, there is a lost opportunity of having been right, and solving the puzzle, as well as being duped by the conspirators; but this loss is not as costly as having pursued a false conspiracy theory, as shown in the table. The following preference order in ordinal ranking $U1 > U4 > U3 > U2$ over the outcomes can then reasonably be assumed (given that there is no other information about the utilities):

Being the first to solve the puzzle

is better than

being misled,

is better than

being naïve, and duped by the conspirators

is better than

being dismissed as a conspiracy theorist in the pejorative sense.

Let us now look at the outcome of applying some standard decision rules to the decision-matrix in table 1. The maximax rule – maximizing the maximum outcome – is based on the decision maker being optimistic about the future. In the puzzle hunt case, the optimal choice under the maximax rule is to believe the conspiracy theory (B_{CT}).¹²

¹² It has been argued that it is difficult to justify the maximax criterion as a rational criterion, as it may seem overly optimistic. The Allais and Ellsberg paradoxes, for example, show that maximizing expectation can lead one to perform

By contrast, the maximin rule is based on the decision-maker being pessimistic about the future. It recommends the decision maker to choose the alternative that has the maximal security level in the sense that it maximizes the minimum payoff. In our case of the puzzle hunt scenario, the matrix for the maximin rule indicates that the appropriate decision is to not believe the conspiracy theory in order to secure the best option out of the worst ones.

4.2 Social power

Social power is a source of non-epistemic value of some conspiracy theories with practical value in exposing injustice or oppression, as well as a way of maintaining or gaining social power. The economist Péter Krekó argues for the importance of groups and social identity in relation to conspiracy belief, in particular how these can drive collective action. He writes that: “Conspiracy theories emerge in groups and are deeply rooted in the social identity of the group. These beliefs help the group and the individual as its member to understand and make sense of the world. They even help draw a line between the in-group and the out-group (see for example “out-group paranoia,” Kramer & Jost, 2002). As these theories are anchored to the social identity of the groups, they can reveal the characteristics of the groups where they emerge, spread, and drive collective action” (Krekó, 2015: 64)

As such, one can utilize belief in conspiracy theories to either prevent, undermine or encourage social upheaval. It can be viewed as an expression of (desired) social power more generally. As van Prooijen and Douglas, (2018) notes, some consequences of believing in conspiracy theories may be positive, depending on the social change that these movements

intuitively sub-optimal actions (See Hájek, 2022). However, as Hanson argues “life would probably be duller if not at least some of us were maximaxers on at least some occasions” (Hansson, 2005, p. 60).

pursue. For instance, conspiracy theories can inspire and justify protest movements¹³. Imhoff and Bruder (2014) consider the role of conspiracy mentality in motivating social action aimed at changing the status quo.

There are many historical examples that can illustrate this. One is Machiavelli (1532/1998) who was very aware of the danger of conspiracies for the nobles. He wrote in *The Prince* that a prince should avoid being hated by the commonfolk, so that he would not have to fear conspiracy. Edward Snowden highlighted the concept of the enemy as fundamental in a conspiracy theory, and that conspiracy theories can “erode civic confidence” (Snowden, 2021: 3). The way Snowden sees it, in democracies today it is not our rights and freedoms that are important, but rather “what *beliefs* are respected: what history, or story, undergirds their identities as citizens, and as members of religious, racial and ethnic communities”. Gosa (2011: 190) likewise discuss the use of conspiracy theories as “counterknowledge” in hip hop culture. Conspiracy theory, as such are an alternative knowledge system in certain subcultures, and they are “intended to challenge mainstream knowledge producers such as the news media and academia”. According to Gosa, conspiracy theories “give voice to inequality”. This does not mean that they always succeed. Rather, as Gosa (2011: 201) notes, wrapping legitimate concerns in conspiracy theories can serve as a digression from the inequality it wishes to expose; and it may hinder “one’s ability to be a powerful cultural force for racial justice”. In this sense there is also a considerable risk one must consider with conspiracy beliefs for social power.

¹³ To be justified here is merely from the view of the individual. I do not discuss the potential further implications of the consequence of such beliefs and movements. The justification refers to the individual taking their actions to be justified.

Consider the following concept to illustrate the social power as a source of non-epistemic value: the patriarchy as a conspiracy. Certain feminist movements characterize the patriarchy in terms of a conspiracy, as the systematic oppression of women by the patriarchy (Hill and Allen, 2021; Haraway, 1988; Knight, 2000). Haraway describes the identification of the other in academic feminism in terms of a conspiracy. She writes: “We have used a lot of toxic ink and trees processed into paper decrying what *they* have meant and how it hurts *us*. The imagined “they” constitute a kind of invisible conspiracy of masculinist scientist and philosophers replete with grants and laboratories” (Haraway, 2003). Further, Knight writes about the role of conspiracy rhetoric in feminism: “In the struggle to give a name to—and find someone to blame for—what Berry Friedan famously called the “problem with no name,” feminists writers have often turned to conspiratorial rhetoric. [...] Metaphors of conspiracy, I want to suggest, have played an important role within a certain trajectory of popular American feminist writing over the last thirty years in its struggle to come to terms with—and come up with terms for—the “problem with no name.”” (Knight, 2000: 117) Knight further discusses the figuration of conspiracy in creating a coherent women’s movement and also how it has become a source of division between feminists. Hill and Allen emphasize the conspiracy framework as a two-edged sword for feminist when identifying the patriarchy. On the one hand “feminist – and others – are utilizing the concept” (Hill and Allen, 2021: 170), and they suggest that it has an important community-forming aspect. But simultaneously, they warn that it can be a “risky strategy for feminists” (165). The reason it is a risky strategy, is because the notion of patriarchy as a conspiracy theory is, as they say, also a key theme in anti-feminist strategy to undermine feminist critiques of power. Further, the authors argue that the discourse that the patriarchy “is ‘nothing more than a conspiracy ‘– created in order to victimize men – has important implications for feminist political claim-making [and] attempts to discredit

feminism.” (184-185) I believe this neatly illustrates how one can utilize conspiracy theories both to expose powerful organizations or people, and also to undermine critique.

Tabel 3 considers alternatives with the perspective of the conspiracy of women being oppressed by the patriarchy. The patriarchy is conspiring to keep women oppressed, and to maintain their power over institutions and other social constructs.

Social power	<i>CT true</i>	<i>CT false</i>
B_{CT}	U1: Undermine patriarchy, no risk.	U2: Undermine patriarchy, risk of backlash
B_{NC}	U3: Continued oppression, missed opportunity	U4: Continued oppression, no missed opportunity

Table 3: Utility preferences of a feminist, to undermine the patriarchy and expose the conspiracy against women in order to stop the oppression.

A reasonable preference relation (again with no information available of probabilities) for the social power cases can be $U1 > U2 > U4 > U3$, as follows:

To undermine the patriarchy with little to no risk

is better than

undermining the patriarchy, with a great risk of backlash and being exposed,

is better than

continued oppression of women, but at least with no lost opportunity of undermining the conspiring powers

is better than

the continued oppression of women with the missed opportunity to expose the conspiring patriarchy.

In the reversed scenario, is the patriarchy undermining the criticism, by utilizing the conspiracy theory, CT^* which says that women are conspiring by falsely accusing the patriarchy of conspiring, labeling it a conspiracy theory. Table 4 represent just such a case. The preference order is the same as in table 3.

Social power	CT^*	$\neg CT^*$
B_{CT}	U1: Undermine feminism, keep power	U2: Undermine feminism, risk of backlash
B_{NC}	U3: Feminism prevails, loss of power and missed opportunity	U4: Feminism prevails, no missed opportunity.

Table 4: Matrix of the patriarchy utilizing CT^ to retain social power, and undermine criticism directed towards it.*

Applying some standard decision rules, maximax and maximin, to the decision-matrix in table 4, starting with the maximax rule in the social power case, the optimal choice under the maximax rule is to believe the conspiracy theory (B_{CT}). On the other hand, considering the maximin rule, based on the decision-maker being pessimistic about the future, the recommendation is choosing the best of the worst alternative, the choice to undermine the patriarchy (alternatively to undermine feminism), with the risk of backlash, U2.

5. Discussion

Arguably, there are other standard decision rules that would be insightful to consider in the cases considered above, such as the regret criterion and the Hurwicz criterion. However, in order to apply these rules, we must deviate from the account above and assign numerical utilities to the alternatives (see table 5 and 6 for the utilities ascribed to the puzzle hunt and social power

criterion respectively). To produce the regret criterion (minimax regret), we assign to each outcome the difference between the utility of the maximal outcome in its column and the utility of the outcome itself (Hanson, 2005). Then the rule recommends to choose the alternative with the minimum of all maximum regrets (see table 7 and 8 for regret matrices for both cases).

Puzzle hunt	CT	$\neg CT$
B_{CT}	50	-10
B_{NC}	-2	10

Social power	CT	$\neg CT$
B_{CT}	50	30
B_{NC}	-10	-2

Table 5 and 6 Numerical utilities for social hunt and puzzle hunt respectively.

Puzzle hunt	CT	$\neg CT$
B_{CT}	0	20
B_{NC}	52	0

Social power	CT	$\neg CT$
B_{CT}	0	0
B_{NC}	60	32

Table 7 and 8 Regret matrices for puzzle hunt and social power respectively.

With social power as non-epistemic utility the minimax regret criterion recommends believing the conspiracy theory (B_{CT}). The same is true for the puzzle hunt case.

A middle way between the pessimism (maximin) criterion and the maximax optimism criterion is the Hurwicz criterion (sometimes called the optimism-pessimism index). According to the Hurwicz rule, we choose the best weighted average payoffs; a compromise between the best

and worst payoffs. We assign the utility values as per table 5 and 6 above and we choose a index, i , between 0 and 1, that reflects the decision-makers level of optimism or pessimism. In the present case I have assigned index .7. A regret matrix may be derived from the above utilities, starting with social power: $(50 \times 0.7) + (30 \times 0.3) = 44$ and $(-10 \times 0.7) + (-2 \times 0.3) = -7.6$, in which case the best choice would be B_{CT} . In the puzzle hunt condition, we get the following values: $(50 \times 0.7) + (-10 \times 0.3) = 32$ and $(-2 \times 0.7) + (10 \times 0.3) = 1.6$, and again the Hurwicz criterion recommends B_{CT} .

One concern commonly raised (as old as decision theory itself!¹⁴) is to what degree the decision maker or the agent can choose what to believe. The discussion so far has been framed in terms believing or not believing in a conspiracy theory being the main options among which the decision makes has to choose. However, nothing hinges on any particular account of belief. Rather, what is required is some form of minimal commitment on the part of the decision maker. Unlike full-blown belief, committing oneself may be under the decision maker's voluntary control. The distinction of acceptance and belief can be useful to consider. For instance, Maher (1993) proposes a decision theoretic account of scientific acceptance of a hypothesis, where accepting a hypothesis involves a voluntary decision to cease inquiring in the matter.¹⁵ Dentith (2021) also discusses the possibility of people sometimes only having a "weak commitment" to a conspiracy theory belief. They argue that some may only be entertaining the idea, and don't necessarily believe it. For example, some conspiracy theorists will claim to be suspicious about politicians or business leaders, leading them to suspect something is not right in the politics. However, as Dentith argues, "a suspicion that something is wrong does not necessarily commit them to any particular conspiracy theory; they may well entertain the idea of a variety of conspiracy theories without necessarily being

¹⁴ As it happens, the first decision theoretic argument in the literature, Pascals' wager concerns doxastic choices.

¹⁵ This does not exclude the possibility to re-evaluate or reassess the question if new evidence or arguments come to light, or new risks may arise.

strongly committed to even one of them. [...] but such suspicions do not entail any strong commitment in a resultant conspiracy theory” (Dentith, 2021: 9902-3).

Further, a decision can be rational without being right and right without being rational. Decision theory as I have engaged with it in this account is concerned with rational decisions, rather than right ones. Instrumental rationality presupposes that the decision maker has some aim, such as to become rich and famous, helping homeless cats, or as in our case to gain social power or be the first to reveal a conspiracy. The aim is external to decision theory, and the recommended choice will be relative to that aim. Relatedly, a reasonable objection to this account might be to question the notion that belief in conspiracy theories can come out as rational on essentially non-epistemic grounds. In response, one may interpret decision theory, as I have indeed intended above, not as a normative theory, but as a descriptive theory of the rational decision maker’s choices, characterising and explaining regularities in the choices that people are disposed to make (Chandler, 2017). Of course, the choices people should make according to rational standard are not completely independent from the choices they in fact make, which is why the normative and the descriptive are not always easy to separate.

However, the subjectivist uncertainty account, as I have argued for here, is an important addition to understanding belief in conspiracy theories. While empirical, and epistemic features of conspiracy theories are one important aspect, I believe a more comprehensive account also acknowledges the social dimension of beliefs in conspiracy theories; a dimension which is difficult to ignore when the phenomenon is so often political, and a moving target, to which sure real-world data are hard to come by. I believe this account has highlighted conspiracy theories as a social phenomenon, and provided an account that looks at the inherently social cognitive dimension of the decision-making process in the beliefs we hold.

6. Conclusion

The philosophical discussion on conspiracy theory has to a large extent centered around the question of whether or not conspiracy theories are justified to believe in, and whether or not we may prima facie dismiss such beliefs as irrational. Empirical research has suggested that most people believe some conspiracy theory. As I have argued, such results beg for an explanation as to what separates rational from non-rational conspiracy belief. I have considered three strategies to this end: the objectivist, the subjectivist under risk and the subjectivist under uncertainty. I have argued that the objectivist framework in its most developed form faces a dilemma: it either grossly overestimate the probabilities of intrinsic failure or provides no guidance as to the viability of a given conspiracy theory. The subjectivist under risk may not be able to provide the necessary probability estimation and quantitative utility assessments. In which case an alternative – or as a compliment to these accounts – I have developed an account of subjectivism under uncertainty about probabilities given non-epistemic utilities. In this model I consider two sources of non-epistemic value, puzzle hunt and social power, showing that whether conspiracy belief is rational depends not only on the type of non-epistemic value, but also on the given decision rule used. Importantly, there are several standard decision rules that do not require any quantitative utility assessments.

Of course, there may exist many more non-epistemic utilities. But my account as it stands, I believe, is an important addition to the conspiracy theory literature in that it, without excluding the epistemic importance of these beliefs, highlights the social aspects of conspiracy theories. The epistemic discussion is one important part of understanding conspiracy theories. But conspiracy theories are infamously moving targets, which makes it difficult to exclusively understand belief in them from an objectivist point of view. This account has argued that much research identifies conspiracy theories as a social phenomenon, which calls for an account that looks at the inherently social cognitive dimension of the decision-making process in the beliefs we hold, especially when facts about the world are not readily available. Utilizing the well-established philosophical tools of decision theory to conspiracy theory theory (the academic study of conspiracy theory), I believe,

may provide helpful nuance to the research and frameworks going forward. Finally, while the account developed here has a distinctive descriptive flavor, it is not without normative import of which the details require a lengthy discussion, which is best left for another paper.

References

- Basham, L. (2018). Joining the conspiracy. *Argumenta*, 3(2), 271-290.
- Brotherton, R. (2015). *Suspicious minds: Why we believe conspiracy theories*. Bloomsbury Publishing.
- Chandler, J. (2017). Descriptive Decision Theory. *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2017/entries/decision-theory-descriptive/>>.
- Coady, D. (2012). *What to believe now: applying epistemology to contemporary issues*. Wiley-Blackwell, Chichester, West Sussex.
- Cassam, Q. (2019). *Conspiracy theories*. John Wiley & Sons.
- Danesi, M. (2002). *The puzzle instinct: the meaning of puzzles in human life*. Indiana University Press; Combined Academic.
- Dentith, M. R. X. (2016). When inferring to a conspiracy might be the best explanation. *Social Epistemology*, 30 (5-6 2016), pp. 572-591, doi: 10.1080/02691728.2016.1172362.
- Dentith, M. R. X. (2019). Conspiracy theories on the basis of the evidence. *Synthese*, 196 (6): 2243-2261.
- Douglas, K.M., Uscinski, J.E., Sutton, R.M., Cichocka, A., Nefes, T., Ang, C.S. & Deravi, F. (2019). Understanding Conspiracy Theories. *Political Psychology*, 40: 3-35. <https://doi.org/10.1111/pops.12568>.
- Doyle, G. (2021). A Bayesian-Rational Framework for Conspiracy Theories.

- Franks, B., Bangerter, A., & Bauer, M. W. (2013). Conspiracy theories as quasi-religious mentality: An integrated account from cognitive science, social representations theory, and frame theory. *Frontiers in Psychology*, 4, Article 424.
- Gosa, T. L. (2011). Counterknowledge, racial paranoia, and the cultic milieu: Decoding hip hop conspiracy theory. *Poetics*, 39(3), 187–204.
- Grimes D.R. (2016). On the Viability of Conspiratorial Beliefs. *PLoS One*, 11(1): e0147905. <https://doi.org/10.1371/journal.pone.0147905>.
- Grimes, D.R. (2021). Medical disinformation and the unviable nature of COVID-19 conspiracy theories. *PLoS One*, Vol. 16, Issue 3, Public Library of Science.
- Hagen, K. (2018). Conspiracy theorists and monological belief systems. *Argumentation*, 3(2), 303-326.
- Hájek, A. (2022). Pascal's Wager, *The Stanford Encyclopedia of Philosophy* (Winter 2022 Edition), Edward N. Zalta & Uri Nodelman (eds.), forthcoming URL = <<https://plato.stanford.edu/archives/win2022/entries/pascal-wager/>>.
- Hansson, SO. (2005). *Decision Theory: A Brief Introduction*. Publisher: Royal Institute of Technology.
- Högström, E. (2014). Leif GW Perssons frågor om Säpospåret. *Expressen*, 26 February 2014. <https://www.expressen.se/nyheter/leif-gw-perssons-fragor-om-saposparet/#:~:text=%22Det%20%C3%A4r%20fiktio%22,teorin%20i%20%22Veckans%20brott%22>.
- Hill, R. L., & Allen, K. (2021). Smash the patriarchy: the changing meanings and work of 'patriarchy' online. *Feminist Theory*, 22(2), 165-189. <https://doi.org/10.1177/1464700120988643>
- Haraway, D. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, Autumn, 1988, Vol. 14, No. 3 (Autumn, 1988), pp. 575-599 Published by: Feminist Studies, Inc.

- Ichino, A., Räikkä, J. (2021) Non-Doxastic Conspiracy Theories. *Argumenta* 7, 1, 247-263.
- Imhoff, R., and Bruder, M. (2014). Speaking (un-)truth to power: Conspiracy mentality as a generalized political attitude. *European Journal of Personality*, 28, 25–43. <https://doi.org/10.1002/per.1930>
- Keeley, B. L. (1999). Of Conspiracy Theories. *The Journal of Philosophy*, 96, 3, pp. 109-126.
- Knight, P. (2000). *Conspiracy Culture: From the Kennedy Assassination to The X-files*. Routledge.
- Krekó, P. (2015). Conspiracy theory as collective motivated cognition. In *The psychology of conspiracy* (pp. 62-76). Routledge.
- Levy, N. (2021). Bad beliefs: Why they happen to good people.
- Lewandowsky, S., Oberauer, K., & Gignac, G. E. (2013). NASA Faked the Moon Landing—Therefore, (Climate) Science Is a Hoax: An Anatomy of the Motivated Rejection of Science. *Psychological Science*, 24(5), 622–633. <https://doi.org/10.1177/0956797612457686>.
- Lewandowsky, S. (2021). Conspiracist cognition: chaos, convenience, and cause for concern. *Journal for Cultural Research*, 25:1, 12-35, DOI: 10.1080/14797585.2021.1886423.
- Machiavelli, N. (1532/1998). *The Prince*, trans. Harvey C. Mansfield, Chicago: University of Chicago Press.
- Napolitano, M. G. (2021). Conspiracy theories and evidential self-insulation. *The Epistemology of Fake News*, 82-106.
- Napolitano, M.G., and Reuter, K. (2021). What is a Conspiracy Theory?. *Erkenntnis*. <https://doi.org/10.1007/s10670-021-00441-6>.
- Olsson, E. J. (2020). Why Bayesian agents polarize. In F. Broncano-Berrocal & J. A. Carter (Eds.), *The Epistemology of Group Disagreement*. Routledge.
- Pigden, C. (1995). Popper revisited, or what is wrong with conspiracy theories? *Philosophy of the Social Sciences* 25 (1):3-34.

- Radnitz, S., and Underwood, P. (2015). Is Belief in Conspiracy Theories Pathological? A Survey Experiment on the Cognitive Roots of Extreme Suspicion. *British Journal of Political Science*, 47(1):1-17, DOI 10.1017/S000712341400055.
- Snowden, E. (2021). <https://www.theguardian.com/commentisfree/2021/jul/01/edward-snowdon-conspiracy-theories-belief-powerlessness>. Accessed on 10 Nov 2022.
- Sunstein, C.R. and Vermeule, A. (2009). Conspiracy Theories: Causes and Cures. *Journal of Political Philosophy*, 17: 202-227. <https://doi.org/10.1111/j.1467-9760.2008.00325.x>.
- Swami, V., Chamorro-Premuzic, T., & Furnham, A. (2010). Unanswered questions: A preliminary investigation of personality and individual difference predictors of 9/11 conspiracist beliefs. *Applied Cognitive Psychology*, 24(6), 749–761. <https://doi.org/10.1002/acp.1583>
- Swami, V., Coles, R., Stieger, S., Pietschnig, J., Furnham, A., Rehim, S., & Voracek, M. (2011). Conspiracist ideation in Britain and Austria: Evidence of a monological belief system and associations between individual psychological differences and real-world and fictitious conspiracy theories. *British Journal of Psychology*, 102(3), 443-463.
- Tsapos, M. (2023). Who is a Conspiracy Theorist? *Social Epistemology*. DOI: 10.1080/02691728.2023.2172695
- van Prooijen, J.-W., & van Vugt, M. (2018). Conspiracy Theories: Evolved Functions and Psychological Mechanisms. *Perspectives on Psychological Science*, 13(6), 770–788. <https://doi.org/10.1177/1745691618774270>.
- van Prooijen, J.-W. and Douglas, K.M. (2018). Belief in conspiracy theories: Basic principles of an emerging research domain. *European Journal of Psychology*, 48: 897-908. <https://doi.org/10.1002/ejsp.2530>.
- Wood, M. J., Douglas, K. M., & Sutton, R. M. (2012). Dead and alive: Beliefs in contradictory conspiracy theories. *Social psychological and personality science*, 3(6), 767-773.