8-1-2024

# Evidence as Dictator: When Arrow’s Impossibility Theorem Meets the Ethics of Belief

Chen-Wei Wu



doi: https://doi.org/10.57709/37137831

# Evidence as Dictator: When Arrow's Impossibility Theorem Meets the Ethics of Belief

by

Chen-Wei Wu

Under the Direction of Neil Van Leeuwen, PhD

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2024

ABSTRACT

Could it be rational to believe contrary to the evidence if the belief brings a substantial amount of practical benefits? In my thesis, I investigate this question through the lens of social choice theory. Specifically, I argue that it is never rational to believe contrary to the evidence by utilizing Arrow's impossibility theorem. To this end, I introduce an analogy between a belief system and a social group, discuss certain conditions that hold in a rational agent's belief system, and compare and analyze the performances of different belief systems. The goal is to shed light on the role of evidence in a rational agent's belief system while exploring the application of theoretical results in social choice theory to the ethics of belief.

INDEX WORDS: Arrow's impossibility theorem, The ethics of belief, Evidentialism, Pragmatism, Dictatorship, Social choice theory

Evidence as Dictator: When Arrow's Impossibility Theorem Meets the Ethics of Belief

by

Chen-Wei Wu

Committee Chair: Neil Van Leeuwen

Committee: Daniel Weiskopf

Electronic Version Approved:

Office of Graduate Services

College of Arts and Sciences

Georgia State University

August 2024

## DEDICATION

This thesis is dedicated to my family, friends, and the Taiwanese community in Atlanta. Their love and support make this work possible.

**ACKNOWLEDGEMENTS**

First of all, I would like to thank my thesis advisor, Neil Van Leeuwen, without whom I wouldn't even start working on this project. His encouragement has been the fuel throughout my journey. And he has always pushed me to think more deeply about the connection between social choice theory and the ethics of belief. I would also like to thank Daniel Weiskopf for his feedback and discussions with me. His comments are critical, sharp, and very helpful. If any of the paragraphs exhibit good quality, they were written when I had his potential response in mind. Many people have given me helpful feedback. I want to thank Armand Babakhanian, Justice Cabantangan, Tony Cheng, Yen-Tung Lee, Hanti Lin, Shawn Standefer, Yun-Cheng Wang, and Mingyan Yang for their feedback on my earlier drafts. Finally, I want to thank Duen-Min Deng for his seminar on decision theory that sparked my interest in social choice theory.

# TABLE OF CONTENTS

**LIST OF TABLES**

# 1 INTRODUCTION

Evidential reasoning is crucial to the rationality of belief. Suppose one sees an Instagram post saying that most COVID-19 vaccines alter one's DNA. Since the claim strikes one as suspicious, one looks for evidence to verify it. After digging into some medical research and reports from fact-checking organizations, one finds out that this claim doesn't conform to the evidence and therefore believes it is false. In mundane cases like this, an agent's belief is rational by virtue of its responsiveness to the evidence.

However, a tradition in epistemology, which can be traced back to William James (1896), advocates that practical reasons also constitute the rationality of belief. Some contemporary followers even propose that, given one's practical reasons, one might be rational to hold a belief, despite the evidence to the contrary (Howard, 2020; Reisner, 2008; Rinard, 2017). To illustrate, consider a variation of the previous example. Suppose the above Instagram post comes from the agent's significant other who staunchly opposes vaccination. Since believing that most COVID-19 vaccines alter one's DNA helps harmonize and consolidate the agent's relationship with her partner, this trend of thought recommends that one should believe it, even though the evidence suggests otherwise.[1]

In this essay, I argue that it is never rational to believe contrary to the evidence by employing Arrow's impossibility theorem, a theorem from social choice theory (Arrow, 1951). To utilize the theorem, my argument builds upon Samir Okasha's (2011) work that connects social choice to theory selection in science. Roughly speaking, the idea is that theory selection in science is relevantly similar to situations in which a society makes collective decisions. Because of their

[1] If consolidating one's relationship with one's significant other doesn't sound like a sufficiently strong practical reason, one can imagine an anti-vaccination institution that will give one a million dollars if one believes that most COVID-19 vaccines alter one's DNA. For a discussion of pragmatism and some other examples, see Worsnip (2021).

similarity, Arrow's theorem, which concerns the latter originally, applies to the former. Expanding on Okasha's analogy, I argue that the theorem equally applies to cases in the ethics of belief, such as the above example.

To give a snapshot of my argument, let's call the collection of mechanisms by which an agent forms and revises cognitive states including beliefs a *belief system*. For the sake of argument, I assume an agent forms and revises beliefs in response to reasons.[2] In other words, reasons influence the outcome of a belief system. Furthermore, I will stipulate that a belief system is rational just in case the agent's beliefs, as well as other cognitive attitudes, are formed and revised in response to reasons in the *right* way, or one of the right ways.[3]

If we analogize a belief system to a social group constituted by different types of reasons, then the formation and revision of beliefs in a belief system is similar to the process through which a society makes collective decisions. With this analogy, Arrow's theorem applies to the processing of a belief system. Informally, the theorem states that a rule via which a society makes a collective decision cannot simultaneously cover all scenarios (unrestricted domain), evaluate an option as preferable when everyone prefers it (weak Pareto), disregard the impact of irrelevant options (independence of irrelevant alternatives), and yield no dictator (non-dictatorship). In parallel, if we label the corresponding conditions of a belief system as (U), (P), (I), and (N), then Arrow's theorem suggests that a belief system cannot simultaneously satisfy all these conditions. In other words, if a belief system satisfies (U), (P), and (I), then it violates (N), the counterpart of the non-dictatorship condition.

[2] In some models of belief formation, beliefs are formed in an unreflective and automatic manner (Egan, 2008; Huebner, 2009; Mandelbaum, 2014). But, even within these models, belief revision can still be reason-responsive.

[3] Just as one might think evidence sometimes permits multiple attitudes toward a proposition (Jackson, 2021), one might also think reasons altogether sometimes permit more than just one attitude. Therefore, I won't assume that there is always a single right way to respond to the reasons.

Having established that the possession of (U), (P), and (I) leads to the violation of (N), the rest of the argument proceeds by suggesting certain properties of a rational belief system. I contend that a rational belief system does satisfy (U), (P), and (I). In addition, I will argue that a belief system in which the evidential reason functions as a dictator performs better than its alternatives. Putting pieces together, the argument of this essay is as below.

Premise 1: If a rational belief system satisfies (U), (P), and (I), then it violates (N).

Premise 2: A rational belief system satisfies (U), (P), and (I).

Premise 3: If a rational belief system violates (N), then the evidential reason should be the dictator in the system.

---

Conclusion: The evidential reason should be the dictator in a rational belief system.

Since a dictator is one whose decisions directly dictate the outcome of a system, it follows from the above conclusion that no other type of reason in a rational belief system can override the decision of the evidence. Therefore, it is never rational to believe contrary to the evidence.

After the introduction, in section II, I will first present a brief exposition of Arrow's theorem and further unpack the analogy between a belief system and a social group. This bolsters the idea that if a rational belief system satisfies (U), (P), and (I), then it violates (N). Next, in section III, I will argue that a rational belief system satisfies (U), (P), and (I). These two claims entail that a rational belief system violates (N). In section IV, I will argue that if, as the violation of (N) indicates, there is one type of reason that functions as a dictator in a rational belief system, then the evidential reason should be the dictator. This will complete my argument. Afterward, in section V, I will address an objection rooted in Amartya Sen's (1970, 1977) response to Arrow's theorem. Finally, I will end by gesturing at future works in the concluding section.

## 2 FROM SOCIAL CHOICE TO BELIEF CHOICE

Let me start with the idea of social choice and some basic formalism. A social choice is a choice among several options made by a group of individuals. For example, since a leader election involves multiple options to choose from and voters jointly make the decisions, a leader election is a social choice. While different societies might utilize different rules for social choice, we can represent all social choice rules as mathematical functions that produce an outcome based on individuals' opinions.

Suppose there is a set of alternatives that a society can choose from, $\{A_1, A_2, \ldots, A_n\}$. The opinion of each individual in the society forms a sequence among the alternatives. As an example, if, in a US presidential election, a US citizen weakly prefers Joe Biden to Donald Trump and weakly prefers Donald Trump to Howie Hawkins, her opinion regarding the three candidates forms the following sequence: Biden ≥ Trump ≥ Hawkins.[4] So understood, the opinions of all voters together constitute a profile of sequences. A social choice rule can be formally represented as a mathematical function that turns a profile of sequences into a single sequence, with the latter representing the group's final preference.

Beyond a mere sequence of alternatives, one might represent social choice rules on top of additional assumptions. For instance, Kenneth Arrow defines a social choice rule as receiving and producing preference orders among the alternatives (1951, p. 23). A sequence of alternatives is an ordering, or a preference order, if and only if the preference relation between alternatives possesses the below two properties (1951, p. 13). For any $A_x$, $A_y$, and $A_z$ in $\{A_1, A_2, \ldots, A_n\}$,

(completeness) $A_x \geq A_y$ or $A_y \geq A_x$, and

[4] In this context, "one weakly prefers $A_1$ to $A_2$" (symbolically, $A_1 \geq A_2$) means that, for the agent, $A_1$ is at least as good as, or as favorable as, $A_2$. From weak preference, we can define strict preference and indifference. Strict preference occurs only when $A_1 \geq A_2$ and $A_2$ is not weakly preferred to $A_1$. Indifference between $A_1$ and $A_2$ exists if and only if $A_1 \geq A_2$ and $A_2 \geq A_1$.

(transitivity) if $A_x \geq A_y$ and $A_y \geq A_z$, then $A_x \geq A_z$.

The merit of making these assumptions is that they enrich the information utilizable for a social choice rule. In this particular case, completeness assumes that every pair of alternatives is comparable, and transitivity prevents cyclic sequences. In other words, they restrict the inputs of social choice rules to not just any sequences of alternatives but "rankings" over alternatives (1951, p. 13).

In place of preference orders, a social choice rule might be represented as receiving a profile of utility functions (Sen, 1970, 1977). A utility function is a function that attaches a real value to each alternative. Consequently, it can represent not only the ordering relation within an individual's opinion among the alternatives but also the specific degree to which one favors an alternative. Nonetheless, since Arrow aims to study how the ordering relation for individuals aggregate, he finds it unnecessary and unjustified to employ such a formalization (1951, p. 17). So, throughout most parts of this essay, I will follow Arrow's setup that defines social choice rules in terms of preference orders. Rules that operate on a profile of utility functions will be discussed in section V.

Social choice rules do not just differ in how they are represented, but also in how desirable they are based on what conditions they satisfy. For example, it is a widespread conviction among social-political thinkers that a society is better off without a dictator (Locke, 1980; Rawls, 2005; Rousseau, 2012; Sen, 1999). To put it in precise terms, an individual is a *dictator* just in case the society to which the individual belongs strictly prefers an alternative to another whenever the individual strictly prefers the former to the latter. With this spirit, a social choice rule is more desirable if it excludes the existence of a dictator.

Surprisingly, however, Arrow's impossibility theorem states that a rational social choice rule is impossible in the sense that the non-dictatorship condition and three other desirable conditions cannot be jointly satisfied. Suppose that a group is choosing among a set of alternatives $\{A_1, A_2, \ldots, A_n\}$ for n larger than two.[5] The theorem states that any social choice rules cannot satisfy all of the following four conditions.[6]

(Unrestricted domain) The domain of a social choice rule is the set of all possible profiles.

(Weak Pareto) If all individuals in a group strictly prefer an alternative $A_x$ to another alternative $A_y$, then the group as a whole strictly prefers $A_x$ to $A_y$.

(Independence of irrelevant alternatives) The choice between $A_x$ and $A_y$ for the whole group can only depend on individuals' preferences for $A_x$ and $A_y$, and not on their preferences over other alternatives.

(Non-dictatorship) There is no individual in the group such that if she strictly prefers $A_x$ to $A_y$, then the group strictly prefers $A_x$ to $A_y$.

For now, I will leave the proofs of the theorem and the discussion of each condition aside.[7] The point of invoking this theorem is that, as the theorem entails, a rule must break the non-dictatorship condition insofar as it satisfies the first three conditions. Consequently, if the analogy between a social group and a belief system is sound, the first premise of my main argument follows directly from Arrow's impossibility theorem.

[5] Some rules, such as the majority rule, can satisfy all four conditions if there are only two alternatives.

[6] To facilitate comprehension, I mostly adopt Okasha's less formal formulations of these conditions (2011, pp. 7-8). The only exception is the formulation of the (independence of irrelevant alternatives). I modify it based on Stegenga's (2015) formulation so as to emphasize that, according to the condition, the choice doesn't depend on individuals' preferences over other irrelevant alternatives. Readers can consult Gaertner (2006, p. 18) for a formal presentation. Also, see Sen (1977, p. 1543) for a version of these conditions spelled out in terms of utility function.

[7] Curious readers can find a nice summary of various proofs for the theorem in the second chapter of Gaertner (2006).

To apply this mathematical result, the analogy between a belief system and a social group is crucial. To make it more intelligible, let me introduce a similar analogy from Okasha (2011). Okasha's analogy links social choice to theory choice in a scientific community. When evaluating competing theories, scientists decide which theory to accept by considering the theoretical virtues of these theories. The situation is akin to a social choice in that they both involve a decision over a set of options and both types of decisions are made by collected individual opinions. One might expect that scientists are compared to voters in Okasha's analogy. However, what Okasha treats as voters in his analogy are theoretical virtues such as accuracy, simplicity, or fruitfulness. Theoretical virtues constrain and, in this sense, act as a voter in the choice of scientific theories. This insight ignites a series of discussions on whether theory choice rules suffer similarly from Arrow's impossibility result and what are available escape routes (Morreau, 2013; Okasha, 2015; Stegenga, 2013, 2015; Weber, 2011). The remaining part of this section examines whether Okasha's insight illuminates the ethics of belief. I argue that it does.

Arrow's impossibility theorem holds in the ethics of belief since the processing of a belief system can be seen as a theory choice at the personal level. Consider again the example in the introduction. When the question of whether most COVID-19 vaccines alter one's DNA is brought to an agent's attention, her belief system encounters several "theories" such as that at least some of the vaccines have such an effect, and that none of the vaccines have such an effect. Moreover, which "theory" an agent eventually believes is determined by factors like how well these "theories" cohere with the evidence and what practical benefits believing such a "theory" brings. So, the inner process of a belief system not only resembles a theory choice in the scientific community but is also similar to a social choice in these crucial aspects. For this reason, I will describe this decision-like process of a belief system as making a *belief choice*.[8]

The outcome of a belief system is determined by various kinds of factors. For example, a strike to the head or consumption of chemical substances might alter one's beliefs. However, the present project aims to depict the interaction between different reasons.[9] More accurately, I aim to study the interaction between reason "types".

The present analogy analogizes a reason type, rather than an individual reason, to a voter.[10] The evidential reason, as a type, indicates the truth of a belief. It favors what is likely to be true. On the other side, the practical reason, as a type, specifies the practical goals facilitated by holding a belief. It tends to endorse a belief that aligns with these practical goals. The present analogy conceives a belief system as a small group constituted by two voters who have different voting tendencies. There are other reason types, but I will omit them to simplify the discussion.[11] Yet, the argument can be generalized to cover cases including more than two reason types.

One might question the validity of this analogy because voters' decisions causally determine a social choice while reason types are causally inert. To reply, Okasha's analogy works even if theoretical virtues do not causally contribute to the selection of a scientific theory. His analogy works because the applicability of Arrow's theorem lies in the shared formal structure

---

[8] To clarify, when I use the word "choice", I don't mean to suggest that an agent chooses to hold a belief *voluntarily* (Alston, 1988; Helton, 2020). What I mean is that the agent undergoes an internal process of narrowing down certain options, which is constrained by some factors.

[9] Here, I will follow Hieronymi (2005) in understanding reasons as considerations. And, I take considerations to be propositions that stand to a belief in an inference-like relation.

[10] Although it is beyond the scope of this essay to discuss the result of adopting an alternative framework that analogizes an individual reason to a voter, I suspect the outcome of this alternative framework is compatible with the present one. But the present framework has a more visible implication on the ethics of belief given that it focuses on reason types.

[11] A potential third party is the moral reason for beliefs. What one is morally obligated to believe might come apart from what one's evidence and practical benefits suggest (Basu, 2019; Gendler, 2011). However, some might be hesitant to think there are moral reasons for beliefs (Sher, 2019). In this essay, I will stay neutral on this issue.

between the two domains. Whether or not Arrow's theorem applies to theory choice in science doesn't hinge on the agency of theoretical virtues. Applying a similar reasoning, Arrow's theorem applies to the ethics of belief regardless of whether reason types are causally inert.

To sum up, I've introduced Arrow's impossibility theorem and presented the analogy between a social group and a belief system. Just like voters in a social group jointly decide their leader, reason types in a belief system jointly decide what an agent believes. Given this analogy, we can formalize the process of making a belief choice in the same way as in social choice theory. This gives rise to Arrow's theorem and, consequently, the first premise of my main argument. In the next section, I will provide the rationale for the second premise of my argument. Together, the first two premises entail that a rational belief system contains a dictator.

## 3 THREE RATIONAL CONSTRAINTS OF A BELIEF SYSTEM

The second premise of my argument states that a rational belief system satisfies three conditions (U), (P), and (I). To facilitate the discussion, let me characterize the process of making a belief choice in more detail. Making a belief choice involves a selection among several "theories." I will call these theories *belief options.* Belief options are theory-like in that they comprise multiple propositions.[12] For example, when an agent confronts a situation that raises consideration about a subject matter, this situation initiates the representation of certain belief options. These belief options are susceptible to reasons available to the agent.[13] Some belief options might be deprived of credence because they go against a huge body of reasons while others are endowed with more cognitive resources and hooked up more closely with one's behavioral repertoire.

To map the above process onto a function receiving a profile of preference orders and returning a ranking among options, consider again the scenario where the agent makes a belief choice regarding whether most COVID-19 vaccines alter one's DNA. In this situation, her belief system evaluates different belief options in light of available reasons. For instance, the information from medical research and fact-checking organizations might favor the belief option that no vaccines can alter one's DNA against the other options, whereas her practical reasons might suggest otherwise. Overall, each type of reason can be represented as having its own preference among the represented belief options. And, eventually, the influences of different reasons result in an unequal distribution of credence among the belief options. This unequal distribution can itself

[12] It is, because of this reason, that most belief choices involve more than just two options. For, typically, there are multiple ways things could go, or in other words, multiple epistemic alternatives involved in a belief choice (Hintikka, 1962; Lewis, 1996).

[13] I emphasize the *availability* of the reasons since belief systems are constrained by an agent's mental and physical conditions (Gigerenzer & Goldstein, 1996; Kahneman, 2003; Simon, 1956).

be represented as a ranking over belief options. Therefore, the whole process that adjudicates between belief options can be formalized as a mathematical function receiving and producing preference orders.

Just like social choice rules differ in their desirability, the process of making a belief choice for different belief systems can vary in terms of their rationality. Let me spell out three conditions that constrain how a rational belief system makes a belief choice. Suppose that a belief system is making a belief choice regarding a set of belief options $\{B_1, B_2, \ldots, B_n\}$, for n larger than two. Each reason type forms a preference order among $\{B_1, B_2, \ldots, B_n\}$. Together, all reason types constitute a profile of preference. A belief system takes the profile and produces a preference order among $\{B_1, B_2, \ldots, B_n\}$ as its outcome. With the above setup, I argue that a rational belief system satisfies the following three conditions.

(U) The input domain of the belief system is the set of all possible profiles.

(P) If all reason types strictly prefer $B_x$ to $B_y$, then the belief system strictly prefers $B_x$ to $B_y$.

(I) The choice between $B_x$ and $B_y$ for the belief system can only depend on the reason types' preferences for $B_x$ and $B_y$, and not on their preferences over other belief options.

In what follows, I will provide some initial reasons for accepting that a rational belief system satisfies these conditions.

What (U) requires is that a belief system produces an outcome for any kind of input. That is, a belief system that satisfies (U) always produces a ranking of belief options no matter how different types of reason rank the options. A desirable election process produces a ranking of candidates regardless of what the voting result is like. Similarly, it seems plausible that a rational belief system produces a ranking of belief options no matter how different reason types rank the options. Hence, a rational belief system satisfies (U).

One might wonder how a belief system satisfying (U) deals with cases in which one allegedly should suspend one's belief. It is advisable to suspend one's belief about a proposition when, for example, all reason types are indifferent between believing and disbelieving a given proposition. If so, doesn't a rational belief system produce no outcome in this case and, therefore, violate (U)?

In response, a belief system that suspends beliefs doesn't violate (U). Even when an agent suspends her belief about a proposition, her belief system could still produce a ranking of belief options. One naive account is that an agent believes all propositions in one of the best options and disbelieves all propositions the negation of which are represented in that belief option. In this case, an agent suspends her belief about a proposition if that belief option does not contain the proposition and its negation. There might be other accounts, but, anyway, a belief system that satisfies (U) can deal with cases of suspended beliefs.

Move on to (P). To simplify the discussion, let us focus on two belief options, $B_1$ and $B_2$. (P) says that if every reason type ranks $B_1$ above $B_2$, a rational belief system should strictly prefer $B_1$ to $B_2$. This constraint might appear intuitive but, to be more rigorous, consider someone whose belief system violates (P). It means that this agent would strictly prefer $B_2$ even though every reason type ranks $B_1$ higher. The only excuse for the agent seems to be that she must have some additional motivations to strictly prefer $B_2$. These motivations can either be reasons or not. If they are individual reasons that are sufficiently strong to change the ranking of some reason types, this would break the assumption that all types of reasons rank $B_1$ higher. However, if not, then the agent stubbornly upholds $B_2$ while ignoring reasons supporting $B_1$. Counting this agent and her belief system as rational is a consequence hard to swallow. Therefore, a rational belief system satisfies (P).

A belief system that satisfies (I) ranks a set of belief options solely based on reason types' preference for these options. More precisely, suppose two input profiles are equivalent with respect to $B_1$ and $B_2$. (I) requires that the resultant rankings of the two input profiles should be the same with respect to $B_1$ and $B_2$.

I contend that (I) is a reasonable constraint on a rational belief system since a rational belief system responds to reasons in one of the right ways. To respond to the reasons rightly is to consider all and only available reasons that are relevant to a belief choice. But reasons are relevant to a choice among a set of belief options only if they are for or against some belief options within this set. To exemplify, consider the question of whether COVID-19 vaccines alter one's DNA. To compare whether it is preferable to believe that some vaccines have the effect or most vaccines have the effect, it seems like one only needs to consider the reasons for or against these belief options, such as the degree to which they are supported by evidence. Whether other belief options are strongly supported by evidence is irrelevant to the comparison of the former two belief options. If so, once a rational belief system has all the relevant reasons for or against a set of belief options determined, then the final ranking of the set of belief options should also be determined.

At this point, one might raise an observation that sometimes considering a new belief option might alter a belief system's choice among a set of options. So, even if the input profile remains the same with respect to a set of belief options, the outcome of a belief system might be different. This idea can be interpreted in two ways. One interpretation is that, after realizing there is a novel alternative $B_4$, the belief system might end up adopting the new belief option instead of $B_1$, $B_2$, and $B_3$. This idea might be plausible when the new option is antecedently unrecognized but convincing. Nonetheless, it is not a counterexample to (I) since it doesn't show that the final ranking for $B_1$, $B_2$, and $B_3$ changes due to considering $B_4$. The belief system might just place $B_4$ ahead of the three belief options without changing the ranking of the latter three.

Another interpretation is that considering a new option does make some options more, or less, appealing than before. One might initially rank a belief option $B_1$ the highest because $B_1$ is the only option among $B_1$, $B_2$, and $B_3$ that explains a phenomenon. However, after recognizing that $B_4$ also explains the phenomenon, $B_1$ is outranked by $B_2$ and $B_3$ since it loses its only advantage. In response, this case is still not a counterexample because the input profiles with respect to $B_1$, $B_2$, and $B_3$ are not the same before and after considering $B_4$. The evidential reason switches from favoring $B_1$ to $B_2$ and $B_3$ to favoring the latter two to $B_1$.

So far, I've offered some preliminary motivations and defenses for thinking that a rational belief system satisfies (U), (P), and (I). These motivations are far from conclusive, but engaging in a thorough discussion requires more space. Hence, I will proceed to explore its consequences. Coupled with a parallel version of Arrow's theorem, the idea that a rational belief system satisfies (U), (P), and (I) entails that it violates the following condition.

> (N) There is no type of reason such that if that type of reason strictly prefers $B_x$ to $B_y$ then the system strictly prefers $B_x$ to $B_y$.

Together, the first two premises of my argument entail that a rational belief system doesn't satisfy (N), which means there is a reason type that serves as a dictator in the system.

This conclusion has substantial implications for the ethics of belief. The existence of a dictator within the realm of reason means that there is a reason type that can "silence" other reason types whenever it strictly prefers a belief option over others. To clarify, a reason type silences another reason type just in case the former dictates the outcome of the belief system so that the latter doesn't play any role in a particular belief choice (Reisner, 2008, p. 21). So, similar to a dictator who has the power to arbitrarily decide the choices of a society, the reason type that functions as a dictator makes choices on behalf of the belief system regardless of other reason types' preferences, when it strictly prefers one belief option to the others.

Conversely, no other reasons can "silence" the reason type that functions as a dictator. If a reason type is silenced, it is rendered irrelevant to a decision. However, since the reason type that functions as a dictator can always directly dictate the outcome of a belief system, it is always relevant to belief choices. Therefore, the reason type that serves as a dictator can never be silenced.

Given that there is a dictator within a rational belief system, it follows that some extant models in the ethics of belief do not accurately depict how a belief system works. For example, Reisner's (2008) and Howard's (2020) models are both not dictatorial in that reason types alternate between the role of a silencer. When the practical stake is high, the practical reason silences the evidential reason. Otherwise, the latter silences the former. For this reason, my argument implies that these models do not correctly depict a rational belief system.

Thus far, I've been concerned about how Arrow's theorem might illuminate the ethics of belief. We've reached the intermediary conclusion that there is a type of reason that functions as a dictator in a rational belief system. Given this result, the question to be addressed in the upcoming section is which type of reason should be the dictator.

## 4 EVIDENTIAL DICTATOR VS. PRACTICAL DICTATOR

Lastly, I argue that if there is a dictator in a rational belief system then the evidential reason should be the dictator.[14] The alternative that will be considered in this section is that the practical reason should be the dictator in a rational belief system. I will refer to a belief system within which the evidential reason functions as a dictator as an *ED system*; otherwise, a *PD system*. The argument in this section assumes that a rational belief system should be practically advantageous. In terms of practical advantage, a system is "better safe than sorry" (Nanay, 2023; Stich, 1990). Based on these assumptions, I will put forth two claims. Firstly, in some crucial cases, an ED system's choice is risk-averse whereas a PD system's choice is risk-taking. Secondly, in these cases, a PD system's choice isn't just risk-taking but unreasonably risky. Since it is "better safe than sorry", a rational belief system should be an ED system..

To see, in what sense, a PD system's choice is risk-taking, let me begin by distinguishing between two types of cases—cases of convergence and divergence. The cases of convergence are those in which both an ED system and a PD system produce the same outcome.[15] To exemplify, mundane belief choices, such as the one presented at the very beginning of this essay, are cases of convergence. With the evidence gathered at hand, an agent equipped with an ED system will believe the claim that none of the COVID-19 vaccines alter one's DNA. Similarly, without a significant other who is against vaccination, an agent with a PD system would hold the same belief

---

[14] Here, I'm using "the dictator" and "a dictator" interchangeably since there can't be more than one dictator in a group. To see this, suppose there are more than one dictator in a group. When the dictators have a conflict of strict preference with regard to two alternatives, $A_1$ and $A_2$, the group will strictly prefer $A_1$ to $A_2$ and also strictly prefer $A_2$ to $A_1$, which is impossible. Therefore, there can only be one dictator in a group.

[15] I will simplify the discussion by considering situations that contain only two belief options, $B_1$ and $B_2$. Under this assumption, the cases of convergence include cases where both the evidential and the practical reason strictly prefer $B_1$ to $B_2$, where one strictly prefers $B_1$ to $B_2$ while another is indifferent between $B_1$ and $B_2$, and where both types of reason are indifferent between $B_1$ and $B_2$.

since believing otherwise can easily lead to trouble. For instance, one might refuse to get vaccinated, rendering one vulnerable to the virus. Generally, belief choices in most ordinary or scientific contexts are cases of convergence since evidentially supported beliefs are often practically beneficial (Howard, 2020; Nanay, 2023; Rinard, 2017). So, to appreciate the crucial difference between an ED system and a PD system, we should zero in the cases of divergence.

The cases of divergence are those in which an ED system and a PD system have divergent outcomes. This happens when the evidential reason and the practical reason have conflicting strict preferences. This is exemplified by the case involving a significant other who is against vaccination. In this case, since the evidence at hand is the same as the case without the vaccine-opposed partner, an agent with an ED system would still believe that none of the vaccines alter one's DNA. By contrast, since this time believing that most of the vaccines have such an effect helps harmonize one's relationship with one's partner, an agent with a PD system would believe it. From cases of divergence like this, we can see a distinctive feature of a PD system, that is, it will sacrifice true beliefs for practical benefits. We can, for instance, substitute the belief in the above example with other beliefs such as that the actor Matthew Perry didn't die in 2023 or even that 1+1=3. No matter how ridiculous or obviously false these beliefs are, an agent with a PD system would believe it so long as the practical benefits it brings are sufficiently large.

Now, the choices of a PD system in cases of divergence are risk-taking in that the agent might be worse off because of the effects of holding false beliefs. As pointed out earlier, believing that most COVID-19 vaccines alter one's DNA can discourage one from getting vaccinated, potentially causing one to be severely ill. A PD system only chooses to form such a false belief because it seems to bring more benefits. In this example, the agent with a PD system might reason that believing otherwise could burn bridges with her partner, a consequence worse than the potential health issues brought by the virus. Hence, it is more beneficial to form such a false belief.

However, I argue that, if we take the below two ideas into account, we should see how the choices of a PD system in cases of divergence, such as the one being discussed, are not just risk-taking but unreasonably risky. The first is the idea that the rationality of a belief system is bounded. Despite being fully rational and going through careful consideration, an agent's practical reasoning can still go wrong because of unexpected relevance. For instance, although the agent foresees how she could get COVID and thinks through how she will take care of herself in that circumstance, she might fail to consider that she can spread the virus to other family members, such as her kids. Though believing that most vaccines alter one's DNA prevents a direct confrontation with the agent's significant other, the cascading effects of this belief, such as the spread of the virus within the family, might end up consuming all the agent's patience for their relationship, which brings their relationship to an end. So, unexpected relevant considerations can change which choices are genuinely more beneficial. Yet, since practical reasoning is constrained by psychological limits, a PD system can fail to consider them and make a risky choice that renders the agent worse off.

Besides the above psychological limits that make practical reasoning imperfect, the second idea is that a good risky decision rests on accurate representations of one's environment. For instance, a financial investment is reasonable, rather than reckless, only if it is based on a good understanding of the market. In the case of belief choice, forming the false belief that most COVID-19 vaccines alter one's DNA is a reasonable choice only if it is true, or at least credible, that holding the belief can sustain her relationship with her significant other. However, a PD system is more likely to lack accurate representations precisely because of its choice in cases of divergence. Given that the agent's belief system is a PD system, her beliefs regarding their relationship are more likely to be a result of flawed practical reasoning. For example, the agent may underestimate the effort she has invested in their relationship and be overly optimistic about her strategy to please her partner. Together, the two ideas illustrate how a PD system's choices in cases of divergence are unreasonably risky. On the one hand, a PD system relies on accurate representations, but, on the other hand, the choices of a PD system compromise the accuracy of representations. These paradoxical features make the choices of a PD system unreasonably risky.

Having established that the choices of a PD system in cases of divergence are unreasonably risky, I will finally sharpen the contrast between the two systems. Indeed, one could agree on the above points but still doubt if an ED system fares better. The concern can be fleshed out in different ways but, in the rest of this section, I will only focus on the following.[16] Doesn't an ED system face a similar paradox? On the one hand, an agent with an ED system also has a bounded rationality. So, the choices of an ED system might be false beliefs. On the other hand, good evidential reasoning equally depends on an accurate representation of one's environment such as what are credible sources of information. If so, why doesn't an ED system's choice in cases of divergence succumb to this paradox, making it just as unreasonable as the choices of a PD system?

In response, I will highlight a relevant difference between the two systems. While an ED system might not be perfectly accurate about one's environment, it strives to achieve it. By contrast, a PD system exacerbates the situation by willingly giving up true beliefs. This is manifested in their reactions to counterevidence. To illustrate, suppose an agent with an ED system falsely believes that most COVID-19 vaccines alter one's DNA because she mistakenly trusts a report from an anti-vaccination institution. After she expresses this particular belief to a medical

[16] There are other worries. For instance, one might argue that practical benefits affect an ED system by shifting the boundary between a belief option being evidentially supported or not (Fantl & McGrath, 2002). Doesn't this effect of practical benefits render an ED system, likewise, vulnerable to the flaws of practical reasoning? Due to the limited space, I will set these worries aside.

expert, she will revise her beliefs or consult more evidence. On the contrary, given the goal of consolidating her relationship with her significant other, an agent with a PD system will ignore what the expert says, if she can, or stay distant from the expert if confronting the expert shatters her belief that most COVID-19 vaccines alter one's DNA. This contrast suggests that the choices of a PD system, in cases of divergence, are unreasonably risky in creating a self-undermining loop, whereas the choices of an ED system aren't.

In conclusion, although a PD system makes similar choices to an ED system in cases of convergence, its choices are risk-taking in cases of divergence. In addition, these risky choices are unreasonable. For this reason, a rational belief system had better be risk-averse, that is, an ED system. In other words, the evidential reason should serve as a dictator in a rational belief system.

To recapitulate things I've discussed thus far, through Arrow's theorem and a comparison between ED and PD systems, I argue that the evidential reason should serve as a dictator in a rational belief system. As an upshot, when stumbling upon a question about what to believe, a rational agent would first consult the available evidence. If her evidence strictly prefers one belief option to another, this fact will settle the question. Only when the evidence is indifferent toward certain belief options, do other types of reason come into play. As long as a belief option is strictly preferred to the others by evidence, a rational agent won't be swayed no matter how large the practical benefits of believing the alternatives might be. Since a rational agent would never hold a belief option that is strictly worse than its alternatives in light of the evidence, it is never rational to believe contrary to the evidence.

## 5 SEN'S ESCAPE ROUTE AND REASONS NOT TO TAKE IT

The discussion in previous sections has proceeded with an Arrovian setup that only considers rules receiving and producing preference orders. However, this setup rests upon assumptions of how a rule should be represented. These assumptions give rise to a challenge. If Arrow's theorem holds only when this Arrovian setup is in place, then one can escape the Arrovian impossibility by rejecting these assumptions (Sen, 1970, 1977). This strategy inspires Okasha (2011) to consider examples in scientific theory selection that break those assumptions and satisfy all four conditions in Arrow's theorem. Making a similar maneuver, one might come up with belief systems that simultaneously satisfy (U), (P), (I), and (N). Such an example would undermine the first premise of my argument. In this section, I consider this challenge and point out why this strategy, perhaps viable in social choice and theory choice, doesn't apply to belief choice.

Let's start by revisiting Sen's response against Arrow's impossibility theorem. As noted earlier in section II, different ways of representing social choice rules vary in the extent to which they presuppose assumptions and exploit information within an individual's opinion. The Arrovian setup employed throughout is often considered informationally parsimonious (Gaertner, 2006, p. 14). The informational parsimony can be unraveled into two aspects, one concerning the measurement of individuals' opinions and another concerning the comparability of different individuals' opinions.[17]

The first aspect in which the Arrovian setup is informationally parsimonious is that it doesn't matter, for an Arrovian social choice rule, how the voters' opinion is gauged as long as the result retains the same ranking. To elucidate the idea, suppose an individual's endorsement of

[17] Although the latter aspect of the Arrovian setup is often labeled as "comparability," I think what it really concerns is the commensurability of different voters' opinions, that is, whether there is a common unit that measures different voters' opinions. See Chang (1997) for the distinction between comparability and commensurability.

social alternatives is quantifiable. For instance, in a situation like a US presidential election, the degree to which a group of voters favors each presidential candidate is measured in one way, and the result is as follows (Table 1).

*Table 1: Voters' Favorability Toward Presidential Candidates (Measure 1)*

| | Trump | Biden | Hawkins |
|---|---|---|---|
| **Voter 1** | 10 | 2 | 1 |
| **Voter 2** | 1 | 3 | 2 |
| **Voter 3** | 1 | 3 | 2 |

Whereas, by another measure, the voters' support is quantified as follows (Table 2).[18]

*Table 2: Voters' Favorability Toward Presidential Candidates (Measure 2)*

| | Trump | Biden | Hawkins |
|---|---|---|---|
| **Voter 1** | 1 | 0.301 | 0 |
| **Voter 2** | 0 | 0.477 | 0.301 |
| **Voter 3** | 0 | 0.477 | 0.301 |

For an Arrovian rule, the change from the values in Table 1 to values in Table 2 is inconsequential since each voter's preference order for candidates remains the same in both cases. Now, consider a rule that takes in a profile of utility functions and operates as below.

*The Utilitarian Rule*: For two alternatives $A_x$ and $A_y$, $A_x$ is weakly preferable to $A_y$ just in case the sum of the values attached to $A_x$ is at least as large as those attached to $A_y$.

If the utilitarian rule is employed as the social choice rule, then the values depicted in Table 1 and Table 2 must be counted as two distinct inputs since they result in different outcomes. This means that, although the two sets of values represent the opinions of the same group of voters, they represent the opinions to different extents and, thereby, vary in their accuracy. But since an

[18] Each value in Table 2 is the logarithm of the corresponding value in Table 1 to base 10.

Arrovian rule cares only about the rankings over the alternatives, it doesn't discriminate between the two sets of values.

Moreover, the second aspect in which the Arrovian setup is informationally parsimonious is that it is also inconsequential, for an Arrovian rule, whether the measure of an individual's opinions is changed independently of others. To compare with an Arrovian rule, consider another rule that doesn't discriminate between the values in Table 1 and those in Table 2.

> *The Maximin Rule*: For two alternatives $A_x$ and $A_y$, $A_x$ is weakly preferable to $A_y$ just in case the value gained by the worst-off individual in $A_x$ is at least as large as the value gained by the worst-off individual in $A_y$.[19]

Like an Arrvoian rule, the maximin rule produces the same outcome for values in Table 1 and Table 2. Nevertheless, it distinguishes the following set of values (Table 3) from the above two.

*Table 3: Voters' Favorability Toward Presidential Candidates (Measure 3)*

| | **Trump** | **Biden** | **Hawkins** |
|---|---|---|---|
| **Voter 1** | 1 | 0.301 | 0 |
| **Voter 2** | 1 | 3 | 2 |
| **Voter 3** | 1 | 3 | 2 |

The result in Table 3 is obtained through the same mathematical function as in the change from Table 1 to Table 2 except that it is done only on the values of one voter. Again, an Arrovian rule would treat the set of values in Table 3 as equivalent to the former two since they constitute the same profile of preference orders.

In contrast, the maximin rule produces a different outcome for Table 3, as opposed to Table 1 and Table 2. Such a change in outcome reflects the fact that the maximin rule involves a comparison of values across different voters. Since specific values for alternatives are compared

[19] This rule is first known as the difference principle in Rawls (2005).

across different voters, the rule presupposes that the values for different voters are measured with a common unit. It follows that, if any mathematical transformations are imposed upon the values, they need to be imposed uniformly across voters. Whereas an Arrovian rule operates on preference orders, no specific values within two different voters will be compared directly. Hence, an Arrovian rule is insensitive to whether a transformation is uniformly imposed across individuals.

To sum up, the Arrovian setup that considers only the rules receiving preference orders is informationally parsimonious because it cares solely about the ranking and involves no comparison of specific values across two voters' opinions. In light of its assumptions in these two respects, the Arrovian setup is referred to as *ordinal non-comparability* (Okasha, 2011; Sen 1977; Stegenga, 2015).

The Arrovian setup excludes some possible social choice rules, such as the utilitarian rule and the maximin rule. As it turns out, it is argued that Arrow's impossibility theorem is an artifact resulting from adopting such an informationally parsimonious setup (Gaertner, 2006). That is to say, some rules exploiting non-ordinal information and allowing interpersonal comparison can satisfy all the four conditions listed in Arrow's theorem. The utilitarian rule is such an example (Sen, 1977, pp. 1545-1546).[20]

Continuing the above line of thought, one might wonder if the utilitarian rule better resembles the processing of a rational belief system than an Arrovian rule. Suppose a belief system is triggered to consider whether most COVID-19 vaccines alter one's DNA. The belief system assesses the degree of support that different types of reason lend to different belief options. Each reason type is represented as a function that assigns values to belief options, with each value

---

[20] The maximin rule can also satisfy the four conditions in Arrow's theorem. The reason I focus on the utilitarian rule, instead of the maximin rule, is that a central concern that motivates the maximin rule is distributive justice (Rawls, 2005, p. 75). While distributive justice provides a reason to adopt the maximin rule in a society, it seems less likely to be a concern for a belief system.

referring to the degree to which that reason type supports a belief option. Let's say the values are as specified below, with None, Some, and Most designating different belief options (Table 4).

*Table 4: Reasons' Favorability Toward Belief Options (Measure 1)*

| | **None** | **Some** | **Most** |
|---|---|---|---|
| **Evidence** | 10 | 2 | 1 |
| **Practicality** | 1 | 3 | 2 |

Mirroring the utilitarian rule, the belief system generates the following preference order: None ≥ Some ≥ Most. Given that the internal process of this belief system resembles the utilitarian rule, it simultaneously satisfies (U), (P), (I), and (N). Therefore, it is a potential counterexample to the first premise of my argument.

The problem with the above illustration of a belief system is that it rests on unsound assumptions. Specifically, I contend that the Arrovian setup is an appropriate framework in the case of belief choice. The internal process of a rational belief system cannot resemble a rule that exploits more than ordinal information or allows interpersonal comparison, such as the utilitarian rule. In the next few paragraphs, I will concentrate on the assumption of interpersonal comparability. The reason is that, as Sen shows, Arrow's theorem can be proven unless the utility functions for different individuals are comparable (1977, pp. 1543-1544). Hence, while relaxing the assumption of ordinal measurement enables us to see the diversity of social choice rules, allowing interpersonal comparison paves the way for escape.

I argue that the assumption of no interpersonal comparison aptly captures the processing of a rational belief system since there is no common unit that measures all reason types. To prompt intuition, different types of reason are like distinct attributes of a person, such as one's wealth and intelligence.[21] Both attributes contribute to, say, the attractiveness of a person. Although there are

[21] This is an example borrowed and modified from Feldman (2020).

numerical measures for both wealth and intelligence, there is no common unit that measures both attributes. For the same reason, we don't typically add one's annual salary and one's result in an IQ test together for an estimate of attractiveness since the two attributes do not contribute to one's attractiveness in an additive manner.

Similarly, assuming that there are numerical measures for each reason type, it doesn't follow that the sum of these values stands for the total support of reasons. Since different reason types target different aspects of belief options, it is questionable to assume that there is a common unit for different reason types and that they jointly support belief options in an additive manner. This idea pinpoints why a rational belief system cannot resemble the utilitarian rule. For the operation of the utilitarian rule involves a summation of the numerical measures for different reason types, but the resultant value might not refer to anything meaningful.

To see the issue more broadly, let us assume otherwise. Suppose there is a common unit for different reason types, and mathematical transformations on the numerical measures for different reason types are always applied uniformly. Each time the numerical measures of the practical reason are modified by adding or subtracting a constant value, then the measures of the evidential reason should change accordingly.

For instance, if we modify the values in Table 4 by subtracting one from the values assigned to each belief option by the practical reason, then the values assigned by the evidential reason should be modified similarly. The result is shown below (Table 5).

*Table 5: Reasons' Favorability Toward Belief Options (Measure 2)*

| | **None** | **Some** | **Most** |
|---|---|---|---|
| **Evidence** | 9 | 1 | 0 |
| **Practicality** | 0 | 2 | 1 |

While such a modification might be benign for the numerical measures of the practical reason, it could substantially alter the meaning for the evidential one. For example, the modification from Table 4 to Table 5 would be benign for the practical reason if its numerical measures track the level of fulfillment of an agent's practical goals. Both 0 and 1, by this measure, could just be ways of indicating "the lowest level of fulfillment". But, as to the evidential reason, one way to measure it has a positive value, a negative value, and zero meaning evidentially favorable, unfavorable, and indifferent, respectively.[22] If this is how the evidential supports are gauged, then the modification isn't appropriate for the evidential reason since a switch from 1 to 0 could suggest a change in what the value stands for.

Generally, different reason types examine different aspects of belief options. As a result, their measures are adopted based on independent considerations. It follows that numerical measures for different reason types could operate on distinct scales. If they operate on distinct scales, some mathematical transformations suitable for one might not be appropriate for the others. For this reason, it should not be the case that mathematical transformations are applied uniformly across the measures of different reason types. This indicates that there is no such common unit for different reason types.

To wrap up, an essential move in Sen's escape route is to abandon Arrow's assumption of no interpersonal comparison. Once the setup allows interpersonal comparison, some rules can satisfy the four conditions in Arrow's theorem. Nevertheless, I've argued that the assumption of no interpersonal comparison aptly captures the processing of a belief system since there is no

[22] A common way to gauge evidential support is to first obtain the ratio of the posterior probability of a proposition to its prior probability and then take its logarithm (Horwich, 1982; Keynes, 1921; Mackie, 1969). With the former value representing the change in the likelihood of that proposition being true, the logarithmic function makes the result easily interpretable since the positivity and negativity of the resultant value correspond to whether the proposition is favorable in light of the evidence (Eells & Fitelson, 2002).

common unit that measures different reason types. Therefore, this strategy doesn’t work in belief choice.

## 6 CONCLUSION

In this essay, I explore a question in the ethics of beliefs by invoking Arrow's impossibility theorem from social choice theory. Along the way, I introduce an analogy between a belief system and a social group in order to leverage the theorem. And then, I argue that a rational belief system satisfies certain conditions. After that, I compare different belief systems to argue that the evidential reason should be the dictator in a rational belief system. Finally, I address a worry that originates from Sen's response to Arrow's theorem. The main takeaway of this essay is that one should listen to one's evidence, despite practical benefits recommending otherwise.

Besides advocating for an epistemological position, I see this essay as a general outline for a larger project. The analogy between the mind and society has been utilized by some philosophers as a lens to inquire into our minds. For example, Plato's analogy between a soul and a city provides insights into the justice, or the right order, of a soul (Ferrari, 2005; Williams, 2001). Following this approach, I believe that the advancement in our knowledge of how society works potentially gives us more resources for studying the mind. That said, there are still lingering questions in the analogy between a belief system and a social group. And, there are more interesting results in social choice theory that awaits exploration. My hope is that this essay could serve as a starting point for future work in this intersection.

**REFERENCES**


Alston, William P. (1988). “The Deontological Conception of Epistemic Justification.” *Philosophical Perspectives*, 2, 257-299.

Arrow, K. J. (1951). *Social Choice and Individual Values*. New York: John Wiley.

Basu, Rima. (2019). “The Wrongs of Racist Beliefs.” *Philosophical Studies*, *176*(9), 2497-2515.

Berker, S (2018). “A Combinatorial Argument Against Practical Reasons for Belief”. *Analytic Philosophy*, 59, 427–470.

Chang, Ruth (1997). “Introduction”. In *Incommensurability, Incomparability, and Practical Reason*. Cambridge, MA, USA: Harvard. pp. 1-34.

Christensen, D. (2021). “The Ineliminability of Epistemic Rationality”. *Philosophy and Phenomenological Research*, 103(3), 501-517.

Eells, E., & Fitelson, B. (2002). “Symmetries and Asymmetries in Evidential Support”. *Philosophical Studies*, 107, 129-142.

Egan, A. (2008). “Seeing and Believing: Perception, Belief Formation and the Divided Mind”. *Philosophical Studies*, 140, 47-63.

Fantl, J., & McGrath, M. (2002). “Evidence, Pragmatics, and Justification”. *The Philosophical Review*, 111(1), 67-94.

Feldman, R. (2000). “The Ethics of Belief”. *Philosophy and Phenomenological Research*, 60(3), 667-695.

Ferrari, G. R. (2005). *City and Soul in Plato’s Republic* (Vol. 2). University of Chicago Press.

Gaertner, W. (2006). *A Primer in Social Choice Theory*. Oxford University Press UK.

Gendler, T. S. (2011). “On the Epistemic Costs of Implicit Bias”. *Philosophical Studies*, 156(1), 33-63.

Gigerenzer, G., & Goldstein, D. G. (1996). “Reasoning the Fast and Frugal Way: Models of Bounded Rationality”. *Psychological review*, 103(4), 650.

Helton, G. (2020). “If You Can’t Change What You Believe, You Don’t Believe It”. *Noûs,* 54(3), 501-526.

Hieronymi, P. (2005). “The Wrong Kind of Reason”. *Journal of Philosophy*, 102 (9), 437-457.

Hintikka, J. (1962). *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Ithaca: Cornell University Press.

Howard, C. (2016). “In Defense of the Wrong Kind of Reason”. *Thought: A Journal of Philosophy*, 5 (1), 53-62.

Howard, C. (2020). “Weighing Epistemic and Practical Reasons for Belief”. *Philosophical Studies*, 177(8), 2227-2243.

Horwich, Paul (1982). *Probability and Evidence*. Cambridge: Cambridge University Press.

Huebner, B. (2009). “Troubles with Stereotypes for Spinozan Minds”. *Philosophy of the Social Sciences*, 39(1), 63-92.

James, W. (1896). “The Will to Believe”. *Philosophical Review*, 6, 88.

Kahneman, D. (2003). “A Perspective on Judgment and Choice: Mapping Bounded Rationality”. *American psychologist*, 58(9), 697.

Keynes, J. M. (1921). *A Treatise on Probability*. London: Macmillan.

Lewis, D. K. (1996). “Elusive Knowledge”. *Australasian Journal of Philosophy*, 74 (4), 549-567.

Locke, J. (1980). *Second Treatise of Government*. 1690. Ed. CB Macpherson. Indianapolis: Hackett.

Mackie, J. L. (1969). “The Relevance Criterion of Confirmation”. *The British Journal for the Philosophy of Science*, 20(1), 27-40.

Mandelbaum, E. (2014). “Thinking is Believing”. *Inquiry*, 57(1), 55-96.

Morreau, M. (2015). “Theory Choice and Social Choice: Kuhn Vindicated”. *Mind*, 124 (493), 239-262.

Nanay, B. (2023). “The Selective Advantage of Representing Correctly”. *Philosophy and Phenomenological Research*, 107(3), 706-717.

Okasha, S. (2011). “Theory Choice and Social Choice: Kuhn versus Arrow”. *Mind*, 120 (477), 83-115.

Okasha, S. (2015). “On Arrow’s Theorem and Scientific Rationality: Reply to Morreau and Stegenga”. *Mind*, 124 (493), 279-294.

Rawls, John (2005). *A Theory of Justice: Original Edition*. Belknap Press.

Rousseau, J. J. (2012). *Rousseau: the basic political writings: discourse on the sciences and the arts, discourse on the origin of inequality, discourse on political economy, on the social contract, The state of war*. Hackett Publishing.

Reisner, A. (2008). “Weighing Pragmatic and Evidential Reasons for Belief”. *Philosophical Studies*, 138, 17-27.

Reisner, A. (2009). “The Possibility of Pragmatic Reasons for Belief and the Wrong Kind of Reasons Problem”. *Philosophical Studies*, 145, 257-272.

Rinard, S. (2017). “No Exception for Belief”. *Philosophy and Phenomenological Research*, 94 (1), 121-143.

Sen, A. (1970). *Collective Choice and Social Welfare*. San Francisco: Holden Day.

Sen, A. (1977). “On Weights and Measures: Informational Constraints in Social Welfare Analysis”. E*conometrica: Journal of the Econometric Society*, 1539-1572.

Sen, A. (1999). *Development as Freedom*. New York: Oxford University Press.

Sher, G. (2019). “A Wild West of the Mind”. *Australasian Journal of Philosophy*, 97(3), 483-496.

Simon, H. A. (1956). “Rational Choice and the Structure of the Environment”. *Psychological review*, 63(2), 129.

Stegenga, J. (2013). “An Impossibility Theorem for Amalgamating Evidence”. *Synthese*, 190 (12), 2391-2411.

Stegenga, J. (2015). “Theory Choice and Social Choice: Okasha versus Sen”. *Mind*, 124 (493), 263-277.

Stich, S. P. (1990). *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation*. The MIT Press.

Weber, M. (2011). “Experimentation versus Theory Choice: A Social-Epistemological Approach”. In Hans Bernhard Schmid, Daniel Sirtes & Marcel Weber (eds.), *Collective Epistemology*. Ontos. pp. 20-203.

Williams, B. (2001). “The Analogy of City and Soul in Plato’s Republic”. E*ssays on Plato’s psychology*, 157-167.

Worsnip, A. (2021). “Can Pragmatists Be Moderate?”. *Philosophy and Phenomenological Research*, 102(3), 531-558.