philosophies

Article

Naturalizing Logic: How Knowledge of Mechanisms Enhances Inductive Inference

Paul Thagard

Department of Philosophy, University of Waterloo, Waterloo, ON N2L 3G1, Canada; pthagard@uwaterloo.ca

Abstract: This paper naturalizes inductive inference by showing how scientific knowledge of real mechanisms provides large benefits to it. I show how knowledge about mechanisms contributes to generalization, inference to the best explanation, causal inference, and reasoning with probabilities. Generalization from some A are B to all A are B is more plausible when a mechanism connects A to B. Inference to the best explanation is strengthened when the explanations are mechanistic and when explanatory hypotheses are themselves mechanistically explained. Causal inference in medical explanation, counterfactual reasoning, and analogy also benefit from mechanistic connections. Mechanisms also help with problems concerning the interpretation, availability, and computation of probabilities.

Keywords: induction; inference; logic; mechanisms; naturalism; probability

check for updates

Citation: Thagard, P. Naturalizing Logic: How Knowledge of Mechanisms Enhances Inductive Inference. *Philosophies* **2021**, *6*, 52. https://doi.org/10.3390/philosophies6020052

Academic Editors: Gordana Dodig-Crnkovic and Marcin J. Schroeder

Received: 17 May 2021
Accepted: 17 June 2021
Published: 21 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

An old philosophy joke (dubiously attributed to Morris Cohen) says that logic texts are divided into two parts: in the first half, on deductive logic, the fallacies are explained; and in the second half, on inductive logic, they are committed. This quip is too hard on inductive inference, which is indispensable in science and everyday life, but it does point to the difference between deduction and induction, which introduces unavoidable uncertainty. This paper argues that an appreciation of mechanisms can help substantially to reduce the problems that attend induction.

Natural philosophy has made substantial progress in integrating epistemology, metaphysics, and ethics with sciences that include physics, psychology, and neuroscience [1]. However, logic might be seen as beyond the reach of naturalism because it provides normative ideals about how people ought to reason, not descriptions of how minds actually work. Nevertheless, advances have been made on the psychology of deduction [2,3], and computational and psychological models of inductive inference have been developed [4,5]. Artificial intelligence has blossomed with applications of deep learning and other kinds of inductive inference [6,7].

This paper explores a different, compatible way of naturalizing inductive inference not just by psychologizing it but also by showing how scientific knowledge of real mechanisms provides large benefits to it. I am not arguing that knowledge of mechanisms is essential to induction, only that some kinds of inductive inference gain substantially from it. Specifically, I show how knowledge about mechanisms contributes to generalization, inference to the best explanation, causal inference, and reasoning with probabilities. Even deduction can be influenced by knowledge about mechanisms in the world.

A Google Scholar search for the term "mechanism" yields more than 7 million responses across fields in the natural and social sciences and more than 200,000 mentions in 2020 alone. Recent philosophy of science has intensely investigated the nature of mechanisms, which can be understood as combinations of connected parts whose interactions produce regular changes [1,8–11]. However, these investigations have neglected the contribution that the understanding of mechanisms makes to the interconnected problems of describing and justifying various kinds of inductive inference. Induction can proceed without mechanisms, but it is more understandable and reliable when inferences are supported by information about mechanisms.

The contribution of mechanisms to good inductive inference is important for practical as well as theoretical reasons. The world is awash in misinformation about issues such as health, climate, and politics. An important way to distinguish information from misinformation is to look at how they differ in their inferential basis. Whereas useful information arises from solid inductive inferences based on mechanisms, misinformation is often based on inferences that either ignore mechanisms or rely on mechanisms that are seriously defective. Hence, an important task for inductive logic is to discriminate strong mechanisms from unreliable ones.

2. Mechanisms

Two of the most pressing problems facing humanity are climate change and viral epidemics. Both problems require inductive inferences to answer important questions, such as whether human industrial activity that produces greenhouse gases is leading to irreversible global warming, and whether new viral diseases such as COVID-19 can be controlled. Fortunately, substantial knowledge has been achieved about the relevant mechanisms as shown in Table 1, which amalgamates the varying terminology used by different philosophers in discussing mechanisms.

Table 1. Mechanisms relevant to global warming and viral epidemics. In the top row, the parentheses show the range of terminology used in philosophical discussions of mechanisms. The other rows describe the operation of climate [12] and viral [13] mechanisms.

	Combination (Whole, System, Structure)	**Parts (Entities, Components)**	**Interactions (Activities, Operations)**	**Changes**	**Results (Behaviors, Functions, Phenomena)**
Global warming	Solar system including Earth	Sun, solar radiation, Earth's atmosphere, Earth's surface, greenhouse gas molecules	Earth absorbs sunlight. Earth emits energy as infrared light. Greenhouse gases absorb light, retaining energy. Energy heats up the Earth.	Earth warms.	Earth's temperature is permanently increasing. Severe weather and flooding are increasingly common.
Viral epidemic	Human population	Bodies, cells, viruses	Viruses infect cells and reproduce. Viruses spread to other bodies.	Infections spread among bodies	Epidemics and pandemics occur.

Consideration of mechanisms should help with two interconnected problems of induction: description and justification. The description problem is to characterize how people typically make inductive inferences. In logic, this would take the form of a set of rules that are applied to premises to generate conclusions. Cognitive science can take a broader view that goes beyond verbal premises and syntax-driven rules of inference to include a variety of mental representations including pictorial and kinesthetic ones and computational procedures different from logical rules of inference. The justification problem asks whether the inductive procedures so described are legitimized by their production of reliable and useful conclusions. I will argue that including mechanisms in the description of several kinds of inductive inference makes them more useful and justifiable.

3. Inductive Generalization

The simplest and most familiar form of induction is generalization from some to all: for example, from the observation that some aardvarks are burrowing animals to the conclusion that all aardvarks burrow. John Stuart Mill noticed that some such inferences are more plausible than others [14], p. 206:

> *Why is a single instance, in some cases, sufficient for a complete induction, while in others myriads of concurring instances, without a single exception known or presumed, go such a little way towards establishing an universal proposition? Whoever can answer this question knows more of the philosophy of logic than the wisest of the ancients, and has solved the problem of Induction.*

For example, a few examples of burrowing aardvarks might suffice to convince us that all aardvarks burrow, whereas we would need many more cases to be confident that all aardvarks eat peanuts.

By "plausible", I mean that a claim is more coherent with the available evidence than opposing claims, although the degree of coherence may not be enough to establish the claim as definitely accepted [15,16]. Here, coherence can be computationally assessed by maximizing the satisfaction of constraints concerning how hypotheses explain evidence and other hypotheses and concerning competition among incompatible hypotheses.

In the 1980s, research in psychology and philosophy provided an answer to Mill's question based on variability [17,18]. Studies found that people who are told that floridium is a metal are quick to infer from a few cases that a high percentage of floridium burns with a blue flame, whereas people who are told that shreebles are birds are much more reluctant to infer that a high percentage of shreebles are blue. A plausible explanation for this difference is that people are aware that metals have little variability in their combustion properties and that colorful birds such as parrots have a lot more variability in their colors. When inductive generalization is about kinds of things and properties that have little variability, then even a single instance or a few of them can suffice.

This analysis probably does capture psychological differences, but a deeper answer comes from considering mechanisms. In the 1780s, Antoine Lavoisier identified the mechanism of combustion as the combination of materials with oxygen to produce heat and light. Much later, the mechanism was deepened to explain how atoms of elements such as carbon interact with atoms of oxygen to produce heat construed as rapidly moving molecules and with light construed as emission of photons. In the floridium induction, we can presume that interaction of the metal atoms with oxygen produces light with a specific frequency, making it easy to infer that floridium burns with a blue flame.

In contrast, the mechanisms in the shreebles case provide much less assurance about the plausibility of the inductive generalization. The main relevant mechanism for inferring the color of animals is genetics on the assumption that offspring inherit color from their parents. However, often, genes do not produce consistent color in species such as cats, parrots, and humans with several different hair and skin colors. Genes have variants called alleles, and different alleles can produce variations in hair color. For example, in humans, most redheads have a mutation in the gene for the melanocortin 1 receptor that affects hair and skin. Accordingly, in the absence of extensive knowledge about the genetics of shreebles, we should be reluctant to infer from a few cases that all or most shreebles are blue.

Mechanisms also help with paradoxes that afflicted attempts in the 1940s and 1950s to establish accounts of inductive generalization as purely syntactic. Confirmation theory proposed that inductive support for generalizations of the form (x)(Fxșș → Gx) came from instances Fa and Ga. For example, observing a black raven confirms the hypothesis that all ravens are black. However, Carl Hempel noticed that (x)(Fx → Gx) is logically equivalent to (x)(~Gx → ~Fx), so it seems that a black raven also confirms the odd hypothesis that all non-black things are non-ravens [19]. Equally oddly, a white shoe confirms the claim that all ravens are black.

The oddity disappears with recognition that hypotheses that connect a kind with a property, as in all ravens are black, are much more plausible when a mechanism connects the kind to the property. Ravens have genes for color that lack alleles for producing colors other than black, and extremely rare white ravens occur because of albinism resulting from mutations in genes for producing the pigment melanin. Hence, the known genetic mechanisms support the confirmation by a black raven that all ravens are black. In contrast, no mechanisms connect non-black things with non-ravens, so we have no reason to take that hypothesis seriously, despite its syntactic equivalence with all ravens are black. One of the lessons of the failure of logical positivism as a philosophy of science is that scientific reasoning is not just a matter of syntax and should instead consider the physical constitution of the world as understood in terms of mechanisms.

A similar resolution is available for Nelson Goodman's new riddle of induction [20]. Examples of green emeralds seem to support the generalization that all emeralds are green, but they also support the generalization that all emeralds are grue, where things are grue if they are observed before time t and green and blue otherwise. Fortunately, mechanisms provide a valuable contrast between the generalizations that all emeralds are green and that all emeralds are grue. The Gem Encyclopia reports (https://www.gia.edu/seeing-green, accessed on 18 June 2021):

Emeralds are formed when chromium, vanadium, and iron are present in the mineral beryl. The varying presence of these three elements gives emerald its range of color. Chromium and vanadium make an intense green color. Iron gives the stone a bluish tint.

The perceived color of emeralds results first from how their constituent elements (parts) interact with light to reflect light in a specific frequency (around 550 nm), and second from how this frequency of light stimulates receptors in the retina to send signals to the brain that get interpreted as green. Nothing in these two mechanisms points to how time t could be relevant to making emeralds blue. As with Hempel's paradox of the ravens, background knowledge about mechanisms is far more useful to understanding inductive generalization than pure syntax.

Inductive generalization assigns a property to a kind, and many philosophers have recognized that natural kinds support induction better than contrived collections [21]. Natural kinds are sometimes assigned metaphysical essences as being the same in all possible worlds, but that account is useless for science-oriented philosophy. A better account of natural kinds was developed by Richard Boyd, who proposed that biological species are clusters of properties held together by underlying properties that are homeostatic: a stable range of properties is maintained because deviations have a low chance of persisting [22]. Hence, we should think of the induction-promoting value of natural kinds as resulting from their underlying mechanisms.

Inductive generalization does not absolutely require knowledge of mechanisms, as sometimes we can have ample instances of A and B to support the conclusion that all A are B even if we lack knowledge of a mechanism connecting A to B. For example, it was known for centuries that willow bark reduces pain before aspirin was isolated in 1897 and its biochemical mechanism was discovered in 1971. Nevertheless, knowledge of mechanisms is highly useful for grasping the contributions of variability and natural kinds to inductive inference and for understanding the failure of the purely syntactic approach of confirmation theory.

4. Inference to the Best Explanation

A narrow use of the term "induction" covers only generalization from some to all, but the broader use covers any inference that differs from deduction in introducing uncertainty. There are many such kinds of induction ranging from analogy to statistical inference, but one of the most common goes by the name "inference to the best explanation" [23–25]. This name was new in the 1960s, but inference to explanatory hypotheses was recognized by nineteenth-century writers such as William Whewell and Charles Peirce, and precursors can be found as far back as Renaissance astronomers and possibly Aristotle. Peirce introduced the term "abduction" for the generation and acceptance of explanatory hypotheses, and much recent work in philosophy and artificial intelligence analyzes varieties of abductive inference [5,26–28].

The basic form of inference to the best explanation is:

Evidence *E* requires explanation.

Hypothesis *H* provides a better explanation of *E* than alternative available explanations.

Therefore, *H*.

This kind of inference, "IBE" for short, is common in everyday life: for example, when people attribute mental states to other people and when mechanics identify causes of automobile breakdowns. It is also common in the law when jurors conclude that an accused criminal is guilty and in medicine when physicians conclude that a patient has a disease that explains the patient's symptoms.

As with generalization, knowledge of mechanisms is not essential to inference to the best explanation, but mechanisms help reduce the inherent riskiness of IBE. The loosest form of IBE has been dismissed as "modus morons":

If *A* then *B*.

B.

Therefore, *A*.

This form of inference is pathetically weak because there may be many other reasons for *B* besides *A*. One way of tightening it up is to require a causal connection, *A* causes *B*, but there may still be other causes that need to considered. By requiring the best explanation, IBE ensures that some comparative assessment of alternatives has taken place.

Advocates of IBE are usually vague about what constitutes an explanation, which most generally is just fitting something puzzling into a familiar pattern. Useful patterns range from the loose storytelling that is frequent in everyday life to the exact deduction found in mathematical fields such as physics. In biology, medicine, cognitive science, and other fields, explanation is often the description of causal mechanism: for example, when influenza is explained by the infection of cells by viruses.

Mechanistic explanations strengthen IBE in two ways. First, mechanisms provide a much tighter connection between hypotheses and evidence than mere if–then relations or abstract causes. The claim that a patient's symptoms of fever, coughing, and pains are the result of influenza can be fleshed out by many causal details, including that a known virus infected the patient's respiratory system, causing specific bodily reactions. When a mechanism is known, we have good reason for taking a hypothesis as a serious contender for explaining symptoms, in comparison with fanciful mechanism-free hypotheses such as that the patient is possessed by demons. IBE still requires that a proposed hypothesis be evaluated according to whether it explains more than alternative hypotheses, but the use of mechanistic explanations sets a high bar for alternatives through the expectation that they should also be able to provide mechanisms that connect the hypothesis with the evidence.

The second way that mechanisms are important to IBE arises because hypotheses are assessed not only by how much they explain but also by the extent to which they themselves are explained [16,29]. For example, in law, the hypothesis that an accused is guilty of murdering a victim has to explain many aspects of the crime scene such as the accused's fingerprints on the murder weapon. However, the guilt hypothesis also gets support by the provision of a motive for why the murderer killed the victim, for example, out of jealousy. Such legal explanations are based on common-sense knowledge that rely on loose psychological mechanisms based on beliefs and desires, for example, the belief that the victim had seduced the accused's spouse, so the accused desired revenge.

In science and medicine, higher-level explanations that explain hypotheses often point to deeper mechanisms supported by substantial evidence. For example, Darwin's theory of evolution gained support from its ability to explain many observations such as the distributions of species, but it was itself explained by the mechanisms of natural selection and the genetic transmission of inherited traits. Skepticism about the truth of scientific theories is inspired by the observation that many scientific hypotheses have turned out to be false: for example, Aristotle's theory of the aether and the chemical theory of phlogiston [30]. However, this pessimistic induction can be countered by the cautiously optimistic induction that all accepted scientific theories that have been deepened by mechanistic explanations have stood up to scrutiny in the face of additional evidence and the competition from alternative theories [31].

The schema for strong IBE is then:

Evidence *E* requires explanation.

Hypothesis *H* provides mechanistic explanations of *E* that are better than alternative available explanations, including alternative mechanisms.

In turn, the mechanisms underlying *H* are explained by more fundamental mechanisms.

Therefore, *H*.

Application of this schema does not completely eliminate the uncertainty of inductive inference but helps to reduce the apparent arbitrariness of IBE. The incorporation of mechanisms helps to overcome the problem identified by Bas van Fraassen that the best explanation might just be the best of a bad lot [32]. If a hypothesis provides a mechanism for the phenomena that constitute the evidence for it, and if this mechanism is itself explained by underlying ones that explain why the parts and their interactions behave as they do, and if both these mechanisms are assessed against alternative explanations, then we have solid grounds for accepting the hypothesis.

The use of mechanisms in support of IBE might seem circular because the existence of mechanisms is itself usually justified by IBE. However, the naturalistic goal is not to provide an a priori justification of inductive inference but rather to identify how induction works when it works well. IBE and inductive inference do not conform to the ideal of deductive inference from indubitable axioms to theorems but require an alternative ideal based on overall coherence among a raft of hypotheses and evidence. Early philosophical ideas about explanatory coherence were vague, but coherence can now be understood mechanistically as a computational process performed by neural networks [33]. The existence of psychologically and neurologically plausible mechanisms for how IBE integrates evidence and hypotheses at multiple levels supports the conclusion that IBE is a good account of much of human induction. Of course, we need to consider alternative hypotheses, and one prominent alternative to IBE is Bayesian inference, as discussed below.

5. Causality and Counterfactuals

One of the most important applications of IBE is to causal claims: for example, the disease COVID-19 is caused by the novel coronavirus; global warming is caused by increasing human emission of greenhouse gases. Such claims go beyond inductive generalizations that all *A* are *B* to assert that *A* causes *B*. These claims are of practical importance as they suggest that we can deal with undesirable effects such as diseases and global warming by modifying their causes. Mechanisms do not explain causality because they presuppose causal notions lightly disguised by saying that the parts and interactions produce, generate, or are responsible for changes.

Analysis of causal inference depends on what causes are taken to be. Skeptics who claim that causality is a bogus, unscientific idea are freed from having to evaluate causal inferences, but they cannot explain why causal talk is ubiquitous in science, as shown by the millions of Google Scholar citations for "cause". David Hume claimed that causality was just constant conjunction [34], which would reduce causal inference to inductive generalization; however, the distinction between correlation and causation is generally acknowledged. Probabilistic theories of causality that look for causes where $P(\textit{effect} \mid \textit{cause}) > P(\textit{effect})$ also try to make causal inference data-driven, but they have trouble with non-observable causes such as subatomic particles. Manipulation theories of causality emphasize how causes can be inferred by identifying interventions that change their effects, but they have trouble with causal relations elsewhere in the galaxy that are beyond human intervention.

I prefer an ecumenical account of causality that avoids definition in favor of identifying standard examples and typical features of causality while noting its explanatory role [1]. Standard examples of cause–effect relations include pushes, pulls, motions, collisions, actions, and diseases. The typical features (looser than necessary and sufficient conditions) of causality are as follows: temporal ordering, with causes before effects; sensory–motor–sensory patterns such as kicking a ball; regularities expressed by general rules; manipulations and interventions; statistical dependencies; and causal networks of influence. Causality explains why events happen and why interventions work.

From this perspective, causality is recognized by inferences to the best explanation that take into account a range of evidence about temporal patterns, correlations, probabilities, and manipulations. Knowledge of mechanisms is not essential to such inferences, but it helps enormously in cases where the interactions of parts connect a putative cause with an effect. For example, the claim that the cause of COVID-19 is infection by the novel coronavirus SARS-CoV-2 is not just correlational, because much is known about how this virus infects cells and disrupts organs such as lungs and blood vessels.

Medical researchers have devoted much attention to analyzing the considerations for inferring the causes of diseases, including the strength of empirical association and background knowledge [35–37]. All of these considerations can be accommodated in computational models based on explanatory coherence [38]. Identifying mechanisms is just one of the considerations that goes into recognizing the coherence of a causal claim, but it provides important backing for claims such as that smoking causes cancer. This hypothesis was accepted in the 1960s before much was known about how cigarette smoke disrupts the normal growth of cells, but it has become all the stronger thanks to understanding of how chemicals are carcinogenic for lung cells. The hypothesis that Zika viruses cause neural defects in infants became more plausible when it was based not just on correlations between Zika infections and birth defects but also on understanding of how the virus infects neurons and produces abnormal growth. In 2021, worries about the occurrence of blood clots in the brains of people who had taken two kinds of vaccines for COVID-19 became more accepted when a mechanism was identified by which the adenovirus-based vaccines cause blood clotting.

Mechanisms are relevant to considering whether a factor *C* is a cause of an event *E* in four situations [39]:

(1) There is a known mechanism by which *C* produces *E*.
(2) There is a plausible mechanism by which *C* produces *E*.
(3) There is no known mechanism by which *C* produces *E*.
(4) There is no plausible mechanism by which *C* produces *E*.

The fourth situation is damning for inductive inference because it suggests that the link between *C* and *E* cannot be given without abandoning well-established science. For example, many paranormal claims such as demonic possession, extrasensory perception, and telekinesis are incompatible with evidence-based physics.

Counterfactuals provide one of the most problematic domains of causal inference. How should we assess such claims as that if the novel coronavirus had not spread to a wet market in Wuhan, then the COVID-19 pandemic would never have happened, or that if the industrial revolution had not occurred, then there would be no global warming? Standard logical treatments of counterfactuals using possible worlds connected by similarity relations are mathematically elegant but scientifically useless.

The AI researcher Judah Pearl developed a much more plausible account of counterfactuals based on causal relations [40]. His account inverts the attempt to analyze causes in terms of counterfactuals: if the cause had not occurred, then the effect would not have happened. Instead, Pearls suggests that counterfactual claims, even though they are not true or false, may yet be plausible or implausible depending on the causal relations in the world. To assess counterfactuals causally, we can work with a causal network and tweak some of the contributory causes to see what happens, either by deleting a cause or by changing the strength of its connection to an effect. Computational methods for such tweaking are available using either Bayesian networks or explanatory coherence networks.

Causal knowledge is not always dependent on mechanisms, but mechanisms enhance causal inference and can also contribute to more plausible counterfactual judgments. Generally, to evaluate a counterfactual claim that if event1 had not happened, then event2 would not have occurred, it helps to ask the following questions. Is there a mechanism connecting event1 to event2? Are there other mechanisms that can produce event2 without event1? For example, consider the counterfactual claim that if Donald Trump had not been infected with the novel coronavirus, then he would not have gotten COVID-19. The mechanisms by which the virus produces the disease are well known, and no other mechanisms produce COVID-19, so the counterfactual about Trump is plausible.

Mechanisms also help with another shaky kind of inductive inference that benefits from causal relations: analogy. At its loosest, analogical inference just notices that two things or events are similar in some respects and infers that they will be similar in another respect. For example, Montreal is similar to Toronto in being a large Canadian city, and Toronto has a subway, so probably Montreal does, too.

Dedre Gentner noticed that analogies are much more useful when they rely on systematic causal relations [41]. If you know the political backgrounds of Canadian cities and how they operate at national, provincial, and municipal levels, then you can construct a causal story about how the decision process that produced a subway in Toronto is likely to have produced a subway in Montreal. Such causal analogies get even stronger from a correspondence between mechanisms operating in the source and target cases. For example, one of the reasons for thinking that the Zika virus causes birth defects is similarity with the mechanism by which measles causes birth defects [38].

For analogical, counterfactual, and causal inference in general, mechanisms are not mandatory. However, they help to reduce uncertainty in inductive inferences that are often error-prone.

6. Probability

Even though probability theory was only invented in the eighteenth century, many philosophers assume that inductive inference should be based on probabilities [42,43]. I find this assumption implausible, because the array of qualitative inferences so far discussed (generalization, IBE, causal, counterfactual, analogical) do not reduce to probabilistic reasoning. Nevertheless, probabilities are indispensable for many kinds of statistical inference: for example, in estimating the effectiveness of vaccines in preventing COVID-19 where data are used to estimate $P\ (infection \mid vaccination)$.

At the core of probabilistic inference is Bayes' theorem, which says that the probability of a hypothesis given the evidence depends on the prior probability of the hypothesis times the probability of the evidence given the hypothesis, all divided by the probability of the evidence. In symbols, $P(H \mid E) = P(H) \times P(E \mid H)/P(E)$. As a theorem of the probability calculus, this result is straightforward, but applying it to real cases of inductive inference faces problems concerning the interpretation, availability, and computation of probabilities. Considerations about mechanisms help with all three of these problems.

The syntax of probability theory is uncontroversial thanks to Kolmogorov's axiomatization, but disputes still rage concerning its semantics [44,45]. Should probabilities be construed as frequencies, degrees of belief, logical relations, or propensities? The frequency interpretation seems most consistent with statistical practices, but it has difficulty in establishing what it means for a probability to be a long-run frequency and in applying this notion to the probability of single events. Bayesians assume that probabilities are degrees of belief but face problems about how such subjective beliefs can objectively describe the world and run up against experimental findings that people's thinking often mangles probabilities [46]. Attempts to describe probabilities as logical relations have encountered problems with describing how abstract considerations of logic and evidence can generate probabilities that satisfy the axioms of probability theory while serving as a practical guide to life.

As a result of these problems, I think the most plausible interpretation of probability is the propensity theory, which says that probabilities are tendencies of physical situations to generate long-term relative frequencies. For example, $P\ (infection \mid vaccination) = x$ is an objective property of the world by which interactions of people, viruses, and vaccines have a disposition to produce over the long run a ratio x of infected people to vaccinated people. However, this interpretation largely ignores the question of the nature of propensities, tendencies, or dispositions.

What does it mean to say that glass has a disposition to break when struck? Fragility is not just a matter of logical relations such as "If the glass is struck, it breaks" or counterfactuals such as "If the glass had been struck, it would have broken." Rather, we can look to the mechanisms by which glass is formed to explain its fragility, including how poorly ordered molecules generate microscopic cracks, scratches, or impurities that become weak points that break when glass is struck or dropped [47]. Similarly, the mechanisms of viral infection, contagion, vaccination, and immunity explain the disposition for people to be protected by vaccines. Mechanisms flesh out the propensity interpretation of probability and point toward a new mechanistic interpretation of probability [1,48].

Karl Popper introduced the interpretation of probabilities as propensities in order to overcome problems faced by the frequency interpretation in applications to single events. He stated that "propensities may be explained as possibilities (or as measures or 'weights' of possibilities) which are endowed with tendencies or dispositions to realize themselves, and which are taken to be responsible for the statistical frequencies with which they will in fact realize themselves in long sequencies of repetitions of an experiment." [49], p. 30.

Similar to forces, propensities point to unobservable dispositional properties of the physical world.

However, Popper did not elucidate the nature of these possibilities, tendencies, or dispositions and did not spell out how they explain frequencies. These gaps are filled by viewing propensities as mechanisms that generate and explain frequencies. Propensities are dispositions to generate frequencies that result from the connections and interactions of the parts in the underlying mechanism. For example, the probability that two dice will roll a total of 12 is 1/36, because the interactions of the dice with their environment and each other will over the long run yield 12 in an approximate proportion of 1/36.

The propensity interpretation, construed mechanistically, works well for statistical probabilities, but it does not apply to the previous kinds of inductive inference considered here. Inductive generalization and inference to the best explanation do not generate conclusions with determinable probabilities because no known propensities generate conclusions such as that all ravens are black and that species evolved by natural selection. Probabilities are just irrelevant to non-statistical judgments [1].

The second problem with Bayesian approaches to inductive inference is that the relevant probabilities are often unavailable, no matter whether they are construed as frequencies, degrees of belief, or propensities. Bayesians usually just present simple examples, but if they worked with examples with tens or hundreds of events or propositions, they would find that Bayesian calculation requires making up vast numbers of conditional probabilities [50]. Paying attention to mechanisms helps to constrain identification of the probabilities that matter in a particular inferential context. For example, understanding the mechanisms for infection, contagion, vaccination, and immunity makes it clear that many extraneous factors can be ignored, such as demonic possession.

Similarly, mechanisms help with the third problem with Bayesian approaches to inductive inference: probabilistic inference has been proven to be computationally intractable in the sense that the amount of computation increases exponentially with the number of variables used [51]. A human brain has thousands or millions of beliefs and large computer data bases can have hundreds or thousands of interrelated variables. Bayesian networks have been developed that prune the potentially explosive networks by introducing a DO operator that makes the networks restricted to causally plausible connections, but the semantics of this operator are ill-specified [52]. Grasping the underlying mechanisms in a situation dramatically prunes the causal relations that provide plausible connections between variables in a Bayesian network, thereby reducing the number of probabilities to be computed.

Thus, knowledge of mechanisms helps with three problems of the Bayesian approach to inductive inference: interpretation, availability, and computation. This assistance is not enough to defend probability theory as a general approach to inductive inference, which would require probabilistic analyses of all the other kinds of induction that I have discussed. However, the legitimate use of probabilities in many important real-life cases of reasoning is enhanced by incorporating knowledge about mechanisms.

7. Evaluating Mechanisms

I have showed the contribution of information about mechanisms to several kinds of inductive inference but have ignored the problem of assessing the quality of mechanisms. To take an extreme example, someone might claim that demonic possession is the mechanism responsible for COVID-19: the parts are demons, organs, and souls, the interactions are that demons invade organs connected to souls, and the results are infected organs and suffering souls. Fortunately, the philosophy of science can assess what makes some mechanisms much more explanatory than others.

Carl Craver and Lindley Darden identify three vices that can occur in representations of mechanisms: superficiality, incompleteness, and incorrectness [9] (ch. 6). Superficial mechanisms merely redescribe the phenomenon to be explained without providing any internal structure, as in Moliere's joke that sedatives put people to sleep because they have dormative virtue. Superficial mechanisms do not seriously compete to be inferred as part of the best explanation of anything. My demon example is not superficial because at least it tries to say something about demons infecting organs and souls to produce symptoms.

Incomplete mechanisms provide only sketches of mechanisms, leaving out crucial parts and interactions. They often have gray or black boxes that need to be filled in. Incompleteness is sometimes unavoidable because of lack of knowledge: for example, when Darwin was unable to explain the inheritance of traits between generations. However, the general aim of science is to fill in the boxes and convert sketches of mechanisms to schemas that provide details about parts, connections, interactions, and causal results. In the evaluation of competing theories, scientists can compare the degree of completeness of the mechanisms they employ in their explanations. My demon example is seriously incomplete because it says nothing about how demons manage to infect bodily organs and how organ changes cause mental suffering.

A mechanism is incorrect if it fails to describe accurately the alleged parts, connections, and interactions, or, equivalently in Craver and Darden's terminology, their entities, organization, and activities. A schema should explain how a mechanism actually works, not just how it might work. At the beginnings of investigation, researchers legitimately can speculate about how a mechanism might possibly work, but evidence should accumulate to suggest that the mechanism is at least plausible in being consistent with background knowledge and ultimately suggest that the mechanism is actually how the world works. We can dismiss my demon mechanism as incorrect because science has found no evidence for the existence of demons and souls, let alone for their activities in causing infections.

In scientific contexts, correctness can be a matter of degree: for example, when good evidence is available for the existence of the parts of the mechanisms but proposed interactions have yet to be empirically established. When alternative mechanisms described by different theories compete, we can compare them with respect to their degree of correctness as well as for their degree of completeness.

An even stronger way in which a mechanism can be incorrect was mentioned above in relation to causal inference. A mechanism that invokes parts and interactions incompatible with legitimate science does not even get to be judged as providing a how-possible explanation. For example, demons have magical capabilities such as taking possession of souls that are incompatible with scientific physics and psychology.

Hence, proposed mechanisms can be evaluated according to their superficiality, completeness, and correctness. We can assess the evidence for the hypothesized parts, connections, and interactions, and for whether the interactions possibly, plausibly, or actually cause the result to be explained.

To sum up the result of this assessment, we can judge a mechanism to be strong, weak, defective, or harmful. A strong mechanism is one with good evidence that its parts, connections, and interactions really do produce the result to be explained. A weak mechanism is one that is not superficial but is missing important details about the proposed parts, connections, interactions, and their effectiveness in producing the result to be explained. Weak mechanisms are not to be dismissed, because they might be the best that can be done currently, as in the early days of investigations of connections between smoking and cancer and between willow bark and pain relief.

More seriously, some mechanisms can be branded as defective because we have good reason to doubt the existence of their parts or interactions, or to doubt the claimed causal production. Doubts about the existence of parts can come from three directions. First, if extensive efforts have failed to find evidence for the parts, then we have reason to believe that they do not exist. The cliché that absence of evidence is not evidence of absence does not apply when extensive attempts to find evidence have failed to provide reason to believe in existence. Extensive attempts to find evidence for demons, unicorns, and gods have provided no evidence for their existence, so belief in non-existence is justified.

Second, hypotheses about proposed parts and interactions can be rejected when the theories that propose them have been superseded by ones that provide better explanations. For example, the phlogiston theory that dominated chemical explanations of combustion for most of the eighteenth century was superseded by Lavoisier's oxygen theory, which proposed different parts and interactions. So, we have good reason to doubt the existence of phlogiston and its interactions with flammable materials. Subsequently, oxygen was isolated from water and other gases, and eventually, oxygen atoms could even be photographed through electron microscopes, so the evidence for the parts and interactions in the oxygen mechanism has become progressively stronger.

Third, the existence of parts and interactions can be doubted because, as already suggested, their operation is inconsistent with established scientific principles. Homeopathic medicine became popular in the early nineteenth century by suggesting that minute quantities of substances that bore some similarity to disease symptoms could be used to cure the disease. This mechanism is defective for explaining diseases because of the general implausibility of causal claims based on minute quantities and similarities.

Some proposed mechanisms are not just defective but actually toxic in that their applications are harmful to human beings by threatening their physical or psychological well-being. The alleged homeopathic mechanism is toxic because people who use ineffective treatments for serious medical problems may fail to get evidence-based treatments that actually work. In the early nineteenth century, homeopathy was probably better than standard treatments based on beliefs about humoral imbalance, which is another defective mechanism lacking in completeness and correctness that was also directly harmful because balance-restoring treatments of bloodletting and purging usually made patients worse.

In summary, mechanisms contribute effectively to inductive inference if they are strong or possibly if they are weak but stronger than available alternatives. Defective and harmful mechanisms block the epistemic and practical effectiveness of inductive inference.

8. Conclusions

The significance of the contributions of mechanisms to inductive inference extends beyond philosophy. Some psychologists have noticed the importance of mechanisms to people's thinking and learning [53,54]. Better understanding of how mechanisms are mentally represented and processed should contribute to further analysis of the advantages of causal mechanisms over relatively superficial knowledge of associations between observable events. Similarly, artificial intelligence has had great successes through the associative inductive method of deep learning, but human-level intelligence will require computers to grasp causality based on mechanisms [6,7]. Both people and computers can learn better if they appreciate the contributions that mechanisms make to inductive inference.

The social significance of the role of mechanisms to inductive inference comes from the need to differentiate misinformation from information. Dealing with climate change and COVID-19 has generated masses of informative evidence, but controversies have also spawned many instances of misinformation such as claims that climate change is a hoax and COVID-19 can be treated by ingesting bleach [55]. Separating information from misinformation requires identifying good patterns of inductive inference that lead to the information and defective patterns that lead to the misinformation. Noting the contribution of mechanisms to justifiable induction is one of the contributions to this separation, as Table 1 summarized for both climate and COVID-19.

It might seem ridiculous that mechanisms could also be relevant to deductive inference, but consider an argument due to Gilbert Harman [29]. Suppose you believe that all aardvarks are gray and that the animal in the zoo is an aardvark. Should you therefore infer deductively that the animal is gray? If you see that the animal is actually brown, you might want to consider instead that it is not an aardvark, or that you were wrong in thinking that all aardvarks are gray. In general, you cannot infer from the premises of a deductive argument that the conclusion is true, because you might need to question some of the premises or even worry about the validity of a particular kind of deductive argument such as disjunctive syllogism. Harman's argument suggests that all deductive inference is actually inductive, so that mechanisms are potentially relevant. You might know, for example, that mutations in color genes are common in aardvark-like animals and thus have further reason to doubt your deductive inference.

I have emphasized that inductive inference does not always depend on mechanisms. Nevertheless, when knowledge of mechanisms is available, it can often be valuable in making inductive inference more reliable. I have shown the relevance of mechanistic information to generalization, inference to the best explanation, causal reasoning, and thinking based on probabilities. Thinking mechanistically makes people smarter and helps to naturalize logic.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Thagard, P. *Natural Philosophy: From Social Brains to Knowledge, Reality, Morality, and Beauty*; Oxford University Press: New York, NY, USA, 2019.
2. Johnson-Laird, P.N.; Byrne, R.M. *Deduction*; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1991.
3. Rips, L.J. *The Psychology of Proof: Deductive Reasoning in Human Thinking*; MIT Press: Cambridge, MA, USA, 1994.
4. Holland, J.H.; Holyoak, K.J.; Nisbett, R.E.; Thagard, P.R. *Induction: Processes of Inference, Learning, and Discovery*; MIT Press: Cambridge, MA, USA, 1986.
5. Thagard, P. *The Cognitive Science of Science: Explanation, Discovery, and Conceptual Change*; MIT Press: Cambridge, MA, USA, 2012.
6. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
7. Thagard, P. *Bots and Beasts: What Makes Machines, Animals, and People Smart?* MIT Press: Cambridge, MA, USA, 2021.
8. Bechtel, W. *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*; Routledge: New York, NY, USA, 2008.
9. Craver, C.F.; Darden, L. *In Search of Mechanisms: Discoveries across the Life Sciences*; University of Chicago Press: Chicago, IL, USA, 2013.
10. Craver, C.F.; Tabery, J. Mechanisms in Science. In *Stanford Encyclopedia of Philosophy*; Stanford University: Stanford, CA, USA, 2015.
11. Glennan, S. *The New Mechanical Philosophy*; Oxford University Press: Oxford, UK, 2017.
12. Ranney, M.A.; Clark, D. Climate change conceptual change: Scientific information can transform attitudes. *Top. Cogn. Sci.* **2016**, *8*, 49–75. [CrossRef] [PubMed]
13. Dimitrov, D.S. Virus entry: Molecular mechanisms and biomedical applications. *Nat. Rev. Microbiol.* **2004**, *2*, 109–122. [CrossRef] [PubMed]
14. Mill, J.S. *A System of Logic*, 8th ed.; Longman: London, UK, 1970.

15. Connell, L.; Keane, M.T. A model of plausibility. *Cogn. Sci.* **2006**, *30*, 95–120. [CrossRef] [PubMed]
16. Thagard, P. Explanatory coherence. *Behav. Brain Sci.* **1989**, *12*, 435–467. [CrossRef]
17. Nisbett, R.E.; Krantz, D.; Jepson, C.; Kunda, Z. The use of statistical heuristics in everyday inductive reasoning. *Psychol. Rev.* **1983**, *90*, 339–363. [CrossRef]
18. Thagard, P.; Nisbett, R.E. Variability and confirmation. *Philos. Studies* **1982**, *42*, 379–394. [CrossRef]
19. Hempel, C.G. *Aspects of Scientific Explanation*; The Free Press: New York, NY, USA, 1965.
20. Goodman, N. *Fact, Fiction and Forecast*, 2nd ed.; Bobbs-Merrill: Indianapolis, IN, USA, 1965.
21. Bird, A.; Tobin, E. Natural Kinds. In *Stanford Encyclopedia of Philosophy*; Stanford University: Stanford, CA, USA, 2017.
22. Boyd, R. Realism, anti-foundationalism and the enthusiasm for natural kinds. *Philos. Studies Int. J. Philos. Anal. Trad.* **1991**, *61*, 127–148. [CrossRef]
23. Harman, G. *Thought*; Princeton University Press: Princeton, NJ, USA, 1973.
24. Lipton, P. *Inference to the Best Explanation*, 2nd ed.; Routledge: London, UK, 2004.
25. Thagard, P. The best explanation: Criteria for theory choice. *J. Philos.* **1978**, *75*, 76–92. [CrossRef]
26. Josephson, J.R.; Josephson, S.G. (Eds.) *Abductive Inference: Computation, Philosophy, Technology*; Cambridge University Press: Cambridge, UK, 1994.
27. Magnani, L. *Abductive Cognition: The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning*; Springer: Berlin, Germany, 2009.
28. Peirce, C.S. *Collected Papers*; Hartshorne, W.P., Burks, A., Eds.; Harvard University Press: Cambridge, MA, USA, 1958.
29. Harman, G. *Change in View: Principles of Reasoning*; MIT Press/Bradford Books: Cambridge, MA, USA, 1986.
30. Laudan, L. A confutation of convergent realism. *Philos. Sci.* **1981**, *48*, 19–49. [CrossRef]
31. Thagard, P. Coherence, truth, and the development of scientific knowledge. *Philos. Sci.* **2007**, *74*, 28–47. [CrossRef]
32. Van Fraassen, B. *The Scientific Image*; Clarendon Press: Oxford, UK, 1980.
33. Thagard, P. *Coherence in Thought and Action*; MIT Press: Cambridge, MA, USA, 2000.
34. Hume, D. *A Treatise of Human Nature*; Selby-Bigge, L.A., Ed.; Clarendon Press: Oxford, UK, 1888.
35. Hill, A.B. The environment and disease: Association or causation? *Proc. R. Soc. Med.* **1965**, *58*, 295–300. [CrossRef] [PubMed]
36. Hennekens, C.H.; Buring, J.E. *Epidemiology in Medicine*; Little, Brown and Co.: Boston, MA, USA, 1987.
37. Dammann, O. *Etiological Explanations*; CRC Press: Boca Raton, FL, USA, 2020.
38. Dammann, O.; Poston, T.; Thagard, P. How do Medical Researchers Make Causal Inferences? In *What Is Scientific Knowledge? An Introduction to Contemporary Epistemology of Science*; McCain, K., Kampourakis, K., Eds.; Routledge: New York, NY, USA, 2019; pp. 33–51.
39. Thagard, P. *How Scientists Explain Disease*; Princeton University Press: Princeton, NJ, USA, 1999.
40. Pearl, J. The algorithmization of counterfactuals. *Ann. Math. Artif. Intell.* **2011**, *61*, 29–39. [CrossRef]
41. Gentner, D. Structure-mapping: A theoretical framework for analogy. *Cogn. Sci.* **1983**, *7*, 155–170. [CrossRef]
42. Howson, C.; Urbach, P. *Scientific Reasoning: The Bayesian Tradition*; Open Court: Lasalle, IL, USA, 1989.
43. Olsson, E. *Against Coherence: Truth, Probability, and Justification*; Oxford University Press: Oxford, UK, 2005.
44. Hájek, A. Interpretations of Probability. In *Stanford Encyclopedia of Philosophy*; Stanford University: Stanford, CA, USA, 2019.
45. Hájek, A.; Hitchcock, C. *The Oxford Handbook of Probability and Philosophy*; Oxford University Press: Oxford, UK, 2016.
46. Kahneman, D.; Tversky, A. (Eds.) *Choices, Values, and Frames*; Cambridge University Press: Cambridge, UK, 2000.
47. Kuhn, J. Part 2: Why does glass break? *Corning Museum of Glass*. 2015. Available online: https://blog.cmog.org/2015/06/03/part-2-why-does-glass-break/ (accessed on 18 June 2021).
48. Abrams, M. Mechanistic probability. *Synthese* **2012**, *187*, 343–375. [CrossRef]
49. Popper, K.R. The propensity interpretation of probability. *Br. J. Philos. Sci.* **1959**, *10*, 25–42. [CrossRef]
50. Thagard, P. Causal inference in legal decision making: Explanatory coherence vs. Bayesian networks. *Appl. Artif. Intell.* **2004**, *18*, 231–249. [CrossRef]
51. Kwisthout, J.; Wareham, T.; van Rooij, I. Bayesian intractability is not an ailment that approximation can cure. *Cogn. Sci.* **2011**, *35*, 779–784. [CrossRef] [PubMed]
52. Pearl, J.; Mackenzie, D. *The Book of Why: The New Science of Cause and Effect*; Basic Books: New York, NY, USA, 2018.
53. Johnson, G.S.; Ahn, W.K. Causal networks or causal Islands? The representation of mechanisms and the transitivity of causal judgment. *Cogn. Sci.* **2015**, *39*, 1468–1503. [CrossRef] [PubMed]
54. Keil, C.F.; Lockhart, K.L. Beyond cause: The development of clockwork cognition. *Curr. Dir. Psychol. Sci.* **2021**, *30*, 167–173. [CrossRef]
55. Thagard, P. Mechanisms of misinformation: Getting COVID-19 wrong and right. Forthcoming.