

# Inference to the Source

Sarwar Ahmed<sup>1\*</sup>

<sup>1\*</sup>Interdisciplinary Centre for Science and Technology Studies (IZWT),  
University of Wuppertal , Gaußstr. 20, Wuppertal, 42119, NRW, Germany.

Corresponding author(s). E-mail(s): [ahmed@uni-wuppertal.de](mailto:ahmed@uni-wuppertal.de);

## Abstract

It has been argued that, in scientific observations, the theory of the observed source should not be involved in the observation process to avoid circular reasoning and ensure reliable inferences. However, the issue of underdetermination of the source has been largely overlooked. I argue that concerns about circularity in inferring the source stem from the hypothetico-deductive (H-D) method. The epistemic threat, if any, arises not from the theory-laden nature of observation but from the underdetermination of the source by the data, since the data could be explained by proposing incompatible sources for it. Overcoming this underdetermination is key to reliably inferring the source. I propose a bidirectional version of inference to the only explanation as a methodological framework that addresses this challenge while circumventing concerns about theory-ladenness. Nevertheless, fully justifying the viability of the background theoretical framework and its accurate description of the source requires a broader conception of evidence. To this end, I argue that integrating meta-empirical assessment into inference to the only explanation offers a promising strategy, extending the concept of evidence in a justifiable manner.

**Keywords:** observation, source, binary black holes, circularity, H-D method, meta-empirical assessments

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Reassessing Methods of Inference to the Source</b>	<b>4</b>
2.1	Hypothetico-deductive Method . . . . .	4
2.2	Data and Phenomena: The Essential Role of Explanation . . . . .	7

2.3	Inference to the Only Explanation: A Two-Step Reasoning . . . . .	10
<b>3</b>	<b>Case study: Observation of a binary black hole Merger</b>	<b>12</b>
3.1	Parameterized post-Einsteinian Framework . . . . .	15
3.2	Limitations of the Parameterized Post-Einsteinian Framework . . . . .	16
<b>4</b>	<b>Back to IOE: Towards a Reliable Inference to the Source</b>	<b>18</b>
4.1	Bidirectional Inference to the Only Explanation: The Case of Binary Black hole Merger . . . . .	18
4.2	Meta-Empirical Support for the Inference to the Source . . . . .	20
<b>5</b>	<b>Conclusion</b>	<b>24</b>

## 1 Introduction

Suppose that theory T predicts x. Suppose further that x is an entity or phenomenon that is imperceptible to our senses but observable by means of our instrumental techniques. Typically, scientists collect representational data through an experimental/instrumental setup that is supposed to represent x. A crucial inferential step in this scientific observation is the inference to the source (x) being observed. Therefore, the source (x) is an observational consequence of the theory T, and the experimental/instrumental setup is designed to detect data originating from it. Note that since modern observations in physics are usually observations of theoretical entities, a theory/model of the source is usually derived from a theory that provides a descriptive account of what the source is and what its basic properties are. In this sense, the source should not be confused with its theory. In other words, the source is a natural phenomenon in the physical world, while the theory of the source is our expected theoretical description. The issue at hand is what is the reliable methodology to infer this natural phenomenon given the detected representational data and our expected theoretical description of it.

Perhaps the most prominent frameworks that come to mind are hypothetico-deductivism(H-D), eliminative reasoning(ER), and inference to the best explanation(IBE). Normally, when one uses eliminative reasoning or IBE, one starts with the evidence, goes through a process of elimination based on the evidence and background knowledge, and ends up with a hypothesis that is accepted as true or provides the best explanation. In contrast, hypothetico-deductive reasoning starts with a hypothesis and, by finding consistent observations for its deduced predictions, ends with a confirmed hypothesis. That is, the direction of eliminative reasoning and IBE is from evidence to hypothesis, whereas the direction of hypothetico-deductive reasoning is from hypothesis to evidence. However, as noted above, what we are dealing with in the process of observation is inference to the source, not confirmation/testing of theory; does the received version of any of these approaches account for this inferential step, and what are the epistemic concerns in relation to each of them?

The philosophical literature is replete with the discussion of theory-laden observation and circularity with regard to the evidential role of observation in testing and confirming theories. However, only a few philosophers have positioned this discussion

within the process of observation itself, that is, theory ladenness and circularity with respect to inference to the source. It has been argued that in order to infer the source reliably, the theory of the source should not be involved in the process of observation (Hacking, 1983; Kosso, 1988; Azzouni, 2004, 1997) or, if it is, there should be an independent empirical access (Franklin, 2002).

These epistemic challenges are not confined to mere theoretical discussions, but are manifested in scientific practice. For example, a number of scientists and philosophers of science have raised concerns about the epistemic aspects of observing binary black hole mergers. Specifically, concerns have been raised about the theory/model ladenness of binary black hole observation and its evidential role in testing general relativity (GR) in this extreme gravitational regime (Yunes and Pretorius, 2009), as well as the alleged circularity of the observations (Elder, 2023).

Since, as will be shown in the body of this paper, the observation of binary black hole mergers does not satisfy any of the normative conditions for inference to the source suggested by the philosophical accounts mentioned, how can one account for it?

In this paper, I argue that concerns about circularity are a consequence of the H-D method. The epistemic threat, if any, comes not from the theory-laden process, but from the underdetermination of the source by information, since the data could be explained by proposing incompatible sources for it. To reliably infer the source is to overcome this underdetermination. I propose a refined method - inference to the only explanation - augmented by meta-empirical assessments to overcome these epistemic challenges.

The paper proceeds as follows:

Section 2 begins by embedding the inferential process of inference to the source in the framework of hypothetico-deductivism, as some philosophers have done, leading to a particular kind of underdetermination that I call “underdetermination of the source by information.” It ends with the conclusion that H-D is not the appropriate method for this inference. In subsection 2.2, I shift the discussion to the relation between data and phenomena, drawing on Massimi (2007), “saving unobservable phenomena,” as a promising starting point for developing a methodological framework for inference to the source. Subsection 2.3 recognizes this bidirectional inference as an eliminative inference, which will be called “inference to the only explanation.” Section 2 concludes with remarks about improvements to this version of eliminative reasoning that make it well suited for the inference to the source in observational processes in scientific practice.

Section 3 is dedicated to presenting a case study that has been the focus of theory-laden observation and alleged circularity, namely the observation of binary black hole mergers. Subsection 3.1 presents the parameterized post-Einsteinian strategy developed to mitigate the theory-ladenness of the observation with respect to testing general relativity in the extreme gravity regime. Subsection 3.2 concludes with remarks on the limitations of these strategies in resolving the “underdetermination of the source by information” thesis articulated throughout the paper.

Section 4 begins with the application of inference to the only explanation method to the case study and recognizes that in order for the method to be epistemically reliable, a broader concept of evidence is essential. Subsection 4.1 and 4.2 show that

both the application of the method and the integration of meta-empirical assessments into it are possible to break the underdetermination in the case of binary black hole mergers. Some concluding remarks are given in Section 5.

## 2 Reassessing Methods of Inference to the Source

### 2.1 Hypothetico-deductive Method

As indicated above, hypothetico-deductive reasoning starts with a hypothesis and, by finding consistent observations for its deduced predictions, ends with a confirmed hypothesis. So, if one embeds inference to the source in this framework, one must normatively reason that the theory of the source being observed should not be involved in the observation process, otherwise it cannot count as a genuine observation. The motivation for this comes from the general understanding of the methodological framework in theory testing and theory confirmation. That is; to truly test a theory, it should not be involved in the observational process. In short, according to this logical reasoning, theory-ladenness leads to a kind of circularity.

If the theory/hypothesis of the source system to be observed is involved in the observational process itself, then the source is underdetermined by the information gathered in the observational process. That is, one cannot reliably infer the source system from the information gathered, which is extracted from data and background knowledge. This epistemic problem affects the reliability of the empirical evidence. In other words, the inferential step appears to be circular in the sense that one must assume the theory of the source to be observed in order to include it in the observational process.

The response of philosophers of science to the epistemic problems I have just outlined is varied. There is a tendency for some philosophers to ignore it as an epistemic problem altogether, focusing more on experimental/instrumental reliability (Shapere, 1982), or in relation to theory testing, some have argued for a case-by-case examination of whether circularity is epistemically threatening (Carrier, 1989; Brown, 1993, 1994). When it comes to circularity regarding inference to the source in an observational process (Hacking, 1983; Kosso, 1988), or inference to the existence of theoretical entities (Azzouni, 2004, 1997) these philosophers have opted for a normative argument that the theory of the source should not be involved in the observational process. Alternatively, it is also argued that an independent test of the theory of the source is required when it is involved in the process of observation (Franklin, 2002). It is also argued that hypothetico-deductive reasoning requires an independent test of a theory that postulates a cause to explain a phenomenon in order to show that the cause is indeed behind the phenomenon (Worrall, 2000). More recently, an alleged circularity has been argued for in the observation of binary black hole mergers (Elder, 2023), which will be the focus of this paper.

In what follows, I provide a more detailed overview of philosophers' responses to the issue of inference to the source in theory-laden observations. Shapere (1982) argues that in many cases of observation, one can identify the theory of the receiver, the theory of transmission, and the theory of the source. According to Shapere, the theory (model) of the source is usually based on well-established available theories

that provide background information. However, on his account, the involvement of the theory of the source being tested in other parts of the observational setting is allowed, as this does not necessarily prevent possible disagreement between prediction and observation results (1982, p. 516, footnote 17). Shapere's central example is the observation of the center of the sun in the case of the solar neutrino experiment where the theory of weak interaction is involved in the theory of the source and the theory of transmission. As it appears, Shapere puts more epistemic load on the theory of the receptor. Namely, if it is reliable, it can detect discrepancies. The reliability, for Shapere, is secured when there is no room for reasonable doubt.

Hacking (1983) describes Shapere's account as "the best extended study of observation," but urges him not to "fall prey to the fallacy of talking about theory without making distinctions" (1983, pp. 183-185). He advances the "argument of independence" as an addition to Shapere's account, and claims that,

Something counts as observing rather than inferring when it satisfies Shapere's minimal criteria, and when the bundle of theories upon which it relies is not intertwined with the facts about the subject matter under investigation (1983, p. 185).

Kosso (1988) by building on Shapere and Hacking, argues that inferring the source of information (x) is reliable only when it is reached through a theory independent of the theory of the source (phenomenon). A similar argument can be found in Azzouni (2004, 1997), who argues that the epistemic credibility of sensory observations derives from their compliance with his four criteria for thick epistemic access. He argues that the epistemic significance of thick epistemic access to theoretical entities derives from the similarity of the robustness of instrumental interaction to that of perceptual observations. That is, like perceptual observations, robust instrumental interactions are independent of the theory of the object (1997, pp. 477-478). Franklin (2002) proposes a list of epistemic strategies for a reliable experiment. One of the strategies is related to inference to the source. Franklin suggests to "use an independently well-corroborated theory of the phenomena to explain the experimental results." According to Franklin, there should be independent empirical support for the theory of the phenomenon.

It follows that, if for a reliable inference to the source, its theory should not be involved in the observational process or there should be independent empirical support for it, and if these are not available, then one cannot reliably infer the source (the reasoning becomes circular).

As we can see, the responses to this epistemic problem revolve around well-established available theories that provide background information, the argument from independence, the argument from robustness, the argument from coincidence, and the argument from consilience. All of these epistemic strategies are valuable in their own right and can be useful in cases where they are applicable.

In the final part of this subsection, I will critically examine the above accounts and address their limitations.

First and foremost, Shapere's account suffers from at least two problems. Firstly, the mere possibility of disagreement between prediction and observation results is far from a justifiable criterion for reliable observation. One should not build reliability on mere possibilities, as this criterion is too weak and lacks epistemic significance. Let's assume H1, the theory (model) under test, represents the source we intend

to infer, while the theory of the receiver is significantly independent of the source theory. According to Shapere's thesis, if the data is consistent with H1, scientists are justified in inferring the source in the manner that H1 represents, which appears permissible. However, what if the theory of the source being tested contributed to the production (processing) of the data? In such a case, if scientists had employed any theory, it would have guided them to infer the source as represented by that theory, since the data would have been consistent with that theory. This situation renders the source permanently undetermined by the data. This is because the raw data, independent from the source theory, cannot conclusively discern the source. The observation of a binary black hole merger by the Laser Interferometer Gravitational-wave Observatory (LIGO) and Virgo collaboration is a case in point since the models played a crucial role in extracting waveforms from the strain data. The strain data on its own not only cannot confirm or disconfirm the theory, it also cannot provide an accurate test for the models which are built based on the fundamental theory since the accuracy of the models is to be assessed in the same process. Secondly, as noted above, the mere reliance on background information obtained from well-established theories makes Shapere's thesis vulnerable to the objection regarding the justifiability of the background information, i.e., no systematic epistemic justification is provided to delimit the possible alternatives that might deviate from the fundamental theory taken as background.

Moreover, one can easily find examples of observations in scientific practice that are based on a single method and principle, and even the theory of the target system is involved in the observation process like the observation of the first binary black hole merger.<sup>1</sup> So, if the proposed criterion for the inference to the source is intended to account for all those cases in scientific practice that are called "observation," then I think it falls short. If not, it is not clear to me how the proposed criterion is supposed to resonate with scientific practice.

In addition to that, a more fundamental problem is the problem of underdetermination. Many modern scientific observations are observations of so-called theoretical entities or processes that are predicted by a particular theory (or particular models of the theory). The properties of these entities or events are represented by the theory that predicts them. Thus, even in cases where the argument from independence holds, one has to give reasons why we should believe that the phenomenon (or the signal, the data) really corresponds to the target system described by the theory or the model and not to something else. In other words, the target system seems to be underdetermined by the evidence, despite the convergence of the samples we collect from different methods of inquiry. One can still argue that even if we are confident that the signal is not an artifact, but rather an original one, the concern remains about determining which target system the signal represents, since different theories or models may predict inconsistent target systems, or the signal might come from something that it is not conceived yet.

All in all, I think the main problem lies elsewhere. As Norton (1994) and Massimi (2004) argue, the argument from underdetermination with all its variants is effective when we take hypothetico-deductivism as our methodological framework. I go further

---

<sup>1</sup>see also Chalmers (2003) for an argument against the argument from independence.

and argue that the problem of theory-ladenness, which underlies all the concerns of Hacking, Kosso, Azzouni, and Franklin, for example, in the case of the inference to the source, arises at least in part from taking hypothetico-deductivism as a methodological framework.<sup>2</sup> This is because, according to this method, the hypothesis to be tested should entail, through its deduced predictions, the evidence that is expected to support it. Therefore, to truly test the hypothesis, it should not be involved in the observational process. This idea is extended to work in the process of observation itself, and has taken the form that the theory of the source of information should not be involved in the process of gathering information about the source that is to be inferred. Worrall (2000, p.66) gives a nice illustration of the difference between demonstrative induction and hypothetico-deductivism. Although he denies that there is a significant difference between the two methods, he points out that hypothetico-deductive reasoning requires an independent test of a theory that postulates a cause to explain a phenomenon, in order to show that the cause is indeed behind the phenomenon. So, whether this epistemic issue is supported by the underdetermination thesis or the theory-ladenness thesis, it can bite when this inferential process of the source system is put into the hypothetico-deductive framework, and I believe that it is not the correct way to reconstruct how inferring the source is really conducted in scientific practice. This should not be taken as a rejection of the H-D framework altogether. Discussions about the justifiability of the H-D method and its applicability in science go far beyond the scope of this paper. The main point to be made here is that this methodological framework cannot account for the inference to the source in observational processes.

The main lesson to be drawn from the above discussion is that philosophical accounts of inference to the source are either too flexible, rendering them epistemically insignificant, or too normative, preventing them from accurately reflecting scientific practice. A more fundamental shortcoming is that the overall inferential framework employed - the H-D method - does not align well with inference to the source in observational processes. In what follows, I will explore an alternative approach that establishes an explanatory link between data and phenomena, ultimately leading to a more robust account of inference to the source: bidirectional inference to the only explanation.

## 2.2 Data and Phenomena: The Essential Role of Explanation

Bas van Fraassen (1980) argues that scientific theories save the observable phenomena, and observable for him is what is perceived by the naked eye. Therefore, according to him, the minimal criteria for accepting a theory is giving a true account of the observable phenomena, which is called empirical adequacy. Later on, van Fraassen (1985, 2001, 2002) relaxes his position and expands the domain of observables to the images that are produced by instruments but remains agnostic regarding any correspondence between these images and the supposed source behind them. It seems his major concern with entities or events that are unobservable to the naked eye was (and still is) that the belief in the instrument-produced phenomena (like electron microscopes and bubble chamber photographs) cannot be extended to the source they

---

<sup>2</sup>see Adam (2004) for a similar reasoning regarding theory testing.

are supposed to represent. This is because, he believes, there is no sensory access to the supposed source to confirm its existence, which implies that sense experience is the only reliable method to observe something.

Arguably one of the most effective critiques of van Fraassen's construal is Bogen and Woodward's (1988) distinction between data and phenomena, where they directly oppose constructive empiricism for failing to recognize this crucial distinction, which blurs the distinction it makes between the observable and the unobservable. That is, according to constructive empiricism, the empirical adequacy of theories derives from saving the phenomena, but the phenomena are not primarily what can be sensorially observed, rather the data are what can be sensorially perceived (1988, pp. 350-351). They further argue that "well-developed theories predict and explain facts about phenomena," whereas data are not "predicted and systematically explained," i.e., data are idiosyncratic and phenomena are robust features of the world (1988, p. 306). According to Bogen and Woodward, data provide evidence for the existence of phenomena, and as long as the evidence and the method of data collection are reliable, we are justified in believing in the "unobservable" phenomena.

Most importantly, Bogen and Woodward argue that it is not only data that provide evidence for phenomena, but facts about phenomena also serve as "evidence for high-level general theories by which they are explained" (1988, p. 306). In other words, general theories are tested against facts about phenomena. Therefore, they disregard the possibility that the general theory is involved in the process of inferring the phenomena. This is because they reject the need of a systematic explanation of the data to assess its reliability.

However, Bogen and Woodward concede that some explanatory connection between data and phenomenon might be necessary to establish a causal connection between them, but they accept it only as a necessary, not a sufficient, condition for the reliability of the evidence, i.e., that E is the evidence for H since other factors like background noise can influence the evidential status of E (1988, p. 341, footnote 34). In their explanation of what kind of explanation this explanatory connection would be, they distinguish two versions of the inference to the best explanation:

First, the notion of "best explanation" is to be understood in the sense that H is the best explanation for E because, given the other alternative causal explanations competing with H, H is the correct one. That is, the alternatives are eliminated based on the truth condition, and then H is taken to be the best explanation for E.

Second, the notion of "best explanation" is to be understood in the sense that H is the best explanation for E because, given the other alternative causal explanations that compete with H, H has the best explanatory power. That is, the alternatives are ruled out because they are not the "best" according to the criterion of explanatory goodness, which in turn is based on a theory of explanation (1988, pp. 338-339).

They correctly point out, I think, that this explanatory link is the first version which I think is better to be called inference to the only explanation. This means, the reliability of the evidence including its explanatory causal connection to the phenomena is established, broadly construed, based on eliminative reasoning. Although throughout the paper Bogen and Woodward reject the possibility of extending the concept of observation to include phenomena as their criteria for observation and



observability is sense perception, the above construal will be helpful to understand the concept of observation in practice including that of binary black holes if one reconstructs this observation in actual scientific practice.

Massimi (2007) picks up the discussion and introduces the role of data models and theoretical models in the process of saving the phenomena. She argues that phenomena appear in data models and data models are “relational structures” which are achieved from the process of data selection and reduction. Theoretical models are various models of the same general theory proposed to save the phenomena. She takes the debate further by giving phenomena a new meaning that is at odds with the empiricist’s understanding of them: instead of images of reality, phenomena are what Kant called objects of experience, and they are what we have epistemic access to (2007, pp. 240-241). Thus, if theoretical models save the phenomena which manifest themselves in data models, and if phenomena include both visually accessible and inaccessible ones, then in the case of having reliable evidence, contrary to constructive empiricism, empirical adequacy extends to visually inaccessible phenomena, and consequently to justifiable belief in theoretical entities proposed by theoretical models. Massimi argues that this can be defended in two ways: first, by showing the essential similarity between the inferential paths to the visually inaccessible phenomena and to the theoretical entity. Second, by presenting a reliable way to select a theoretical model that best fits the data model (2007, p. 249).

She continues taking the second way as follows. A background theory consisting of a family of models imposes constraints on the possible expected values for the parameter that the data model is intended to measure. Thus, the background theory helps to provide information in this inference process. But the actual probability calculation of the intended parameter is done by incorporating the properties of the theoretical entities suggested by the theoretical models and the value of the phenomenon manifested in the data model. Since multiple theoretical models that may refer to different theoretical entities may compete to explain the observed phenomenon in the data model, the model that best fits the phenomenon given the empirical and theoretical constraints is selected. That is, the theoretical entity manifests itself in a bidirectional way, the “upward path” from the data model (the experimental result) and the “downward path” from the theoretical model (the expected parameter value), and eventually they converge to save the phenomenon (2007, pp. 249-250).<sup>3</sup> Note that this inference process that Massimi presents here is a more sophisticated version but basically similar to what Bogen and Woodward (1988) present as the first version of inference to the best explanation, which I call inference to the only explanation.

I should highlight a few reasons to show why I consider IOE to be a type of eliminative reasoning and to distinguish IOE from IBE. First, as noted above, IOE does not rely on explanatory power to rank the candidates and select the best explanation, but relies on the truth condition and rejects those alternatives that conflict with the evidence (as noted, both theoretical and empirical evidence). Second, it does not have a comparative nature to select the “best” explanation, but it necessarily eliminates all but one, which is why it is better understood as eliminative reasoning. Third, IBE

---

<sup>3</sup>Wallace (1992) also argues that Galileo used a bidirectional inference in his observations.

standard reasoning does not have to deal with the problem of unconceived alternatives, i.e., the set of considered explanations need not be exhaustive, especially when the IBE mode of inference is used to confirm the viability of a statement rather than its truth. However, as I reconstruct the methodology from the scientific literature on the observation of binary black holes, bidirectional IOE must address the problem of unconceived alternatives. Fourth, IOE could be understood as an instance of IBE if the set of considered alternatives contains only one candidate and no eliminative reasoning is involved (Bird, 2007), but as we will observe in the our case study, the collaboration employs systematic eliminative reasoning to exclude, for example, binary neutron stars and other scenarios.

In the following section I build on Massimi and Bogen and Woodward to introduce inference to the only explanation in the sense that can be applicable, unlike what they have done, to the actual process of the inference to the source in observational situations in scientific practice.

### 2.3 Inference to the Only Explanation: A Two-Step Reasoning

Following Dorling (1973), Bogen and Woodward (1988), Bird (2005, 2007, 2010) and Woodward (2024), I take eliminative reasoning as explanatory reasoning and call it inference to the only explanation.<sup>4</sup> This is because the use of the concept of explanation works well when it comes to the inference to the source of information in an observational process, since it has an informational virtue and informs us about the source.<sup>5</sup> The essential difference between IBE and IOE is that the latter, contrary to the former, does not rely on the explanatory power of the hypothesis during the selection process among the alternatives. Instead, it relies on the evidence and theoretical background knowledge. Therefore, the inferential process of IOE is eliminative, using the evidence and background knowledge to eliminate all alternatives but one by considering them false. More specifically, in IOE, the best explanation is considered to be the best explanation because all alternatives are considered not to be explanations based on the evidence and background knowledge.

Probably the best way to illustrate IOE is a two-step argument: First, delimit the space of alternatives; second, systematically eliminate all but one of the alternatives based on evidence (evidence being understood as a combination of empirical and theoretical evidence). The two-step strategy is well documented in the literature (Kitcher, 1993; Norton, 1995; Bird, 2005, 2010; Forber, 2011; McCoy, 2021; Woodward, 2024). For instance, Norton states eliminative reasoning as follows.

I shall construe eliminative inductions broadly as arguments with premises of two types: (a) premises that define a universe of theories or hypotheses, one of which is posited as true; and (b) premises that enable the elimination of members of this universe by either deductive or inductive inference (1995, p. 30).

For Norton, the second step is based on observations, principles, laws, and well-established theories which have an empirical basis. As for the first step, it is based on those assumptions and arguments that are general enough and uncontroversial enough

---

<sup>4</sup>In the literature, it is also called demonstrative induction, eliminative induction, deduction from phenomena, Holmesian inference, and eliminative inference.

<sup>5</sup>Similar reasoning can be found in Salmon (2001), Cabrera (2017) and Woodward (2024).

to encompass a large number of alternatives. For example, he argues that this kind of reasoning is applicable to Einstein's discovery of general relativity. In the first step, Einstein assumed that the gravitational tensor is proportional to the stress energy, and from this he identified the possible set of field equations. In the second step, based on a number of principles such as general covariance, equivalence, conservation of energy-momentum, and the corresponding Newtonian and special relativistic limit, he eliminated all possible field equations and saved Einstein's field equations (1995, p. 31).<sup>6</sup> Norton explicitly argues that the inductive risk in eliminative reasoning is shifted to be located in the premises of greater generality, that is in those premises that delimit the number of possible explanations. Furthermore, he argues that the premises of greater generality become points of doubt only during major scientific revolutions (1994, p. 20). Massimi argues that "the conclusion of a demonstrative induction is empirically adequate with respect to the phenomena listed in the phenomenal premises only insofar as it is also theoretically adequate with respect to the theoretical constraints of the major premises" (2004, p. 269). In Massimi's terms, phenomenal premises and major premises denote premises of lesser generality and premises of greater generality in Norton's terms, respectively. In other words, she argues that the conclusion of demonstrative induction is adequate in so far as both steps of the eliminative inference, delimiting the possible alternatives and eliminating the alternatives but one, are justified. However, she does not give an account of where the justifiability of the theoretical constraints comes from. I will come back to this point later.

Now, how can one incorporate inference to the only explanation into the observational setting itself? Recall that Massimi (2007) calls her version of inference to the only explanation "saving the unobservable phenomena," that is, she abides by the empiricist tradition that links observation to perception. I think this reasoning, combined with some improvements, is the correct way to reconstruct how the source is inferred in observational processes in scientific practice. As we have seen, this specific version of the IOE in Massimi (2007) has preserved its two-step reasoning; the only difference that can be recognized is that the elimination process and the selection of the best candidate in the second step are bidirectional. That is, the delimitation of the possible explanations (in the form of theoretical models which introduce theoretical entities) based on background knowledge, which she calls theoretical constraints, is the first step in the inference to the only explanation. Without constraining the possible explanations, the bidirectional elimination process is not possible. However, as in her 2004 paper as well, Massimi does not provide an epistemic justification for the process of constraining the possible explanations based on theoretical constraints.

Before addressing these problems and finally applying inference to the only explanation presented here as inference to the source in observational processes, I present a case study of scientific practice. An observation in scientific practice that seems to share the feature discussed so far, namely the involvement of the theory of the source in the observational process, which, according to the hypothetico-deductive method, leads to circularity.

---

<sup>6</sup>Norton has identified the same reasoning in Bohr's atomic theory, Newton's law of gravitational attraction, and blackbody radiation see (Norton, 1987, 1994, 2000).

### 3 Case study: Observation of a binary black hole Merger

On the 11th of February 2016 the LIGO scientific collaboration and Virgo collaboration published a paper in which they report that on September 14, 2015 the two detectors of the Laser Interferometer Gravitational-Wave Observatory simultaneously observed a transient gravitational-wave signal. From that they announce that “this is the first direct detection of gravitational waves and the first observation of a binary black hole merger”(Abbott et al., 2016). This was followed by a storm of other observations of black holes and neutron stars. These observations opened a new window for observing the previously unobservable parts of the universe and expanded the spectrum of the observing channels through which we observe the external world. It is worth mention that in this paper I am predominantly concerned with the first announcement of the observation of binary black hole merger as an accepted observational claim in a specific period of time. Therefore, my discussion will for the most part revolve around the observation of the binary and consider the gravitational waves as an information channel.

The existence of gravitational waves is a prediction of general relativity as Einstein developed the idea of wave solutions for his linearized weak-field equations. These waves were conceived as spatial strains that travel at the speed of light.

The wave-form solutions to Einstein’s field equations describe gravitational waves originating from a source, such as a compact binary, and propagating through gravitational potential (curvature perturbation). However, Einstein’s field equations do not provide exact solutions for compact binaries (Yunes and Pretorius, 2009; Kennefick, 2007; Patton, 2020; Elder, 2023). This is known as the general relativistic two-body problem.

Scientists have addressed this problem by developing approximation schemes derived from idealized solutions of the field equations to model compact binaries. These approximations have allowed them to predict the gravitational waves that compact binaries may produce. These models fit into Patrick Suppes’s categorization of models, providing experimentally testable hypotheses (testable predictions) for a specific physical phenomenon embedded in the background theory (in this case, general relativity). The accuracy of these models varies depending on the nature of the background theory (Suppes, 1960, 2009). This is also how Massimi (2007) understands the role of theoretical models in determining the source.

As previously noted, the phenomenon in question, namely the merging process of two orbiting compact objects, requires several models within the framework of general relativity to be adequately represented. For simplicity, the merging evolution of compact binaries is divided into three phases: inspiralling, merger, and ring-down. Therefore, different models are developed to describe these phases due to significant differences in their physical characteristics.

Although the empirical signal, which can be counted as a data model, already indicates some characteristics of the source, without these theoretical models, which

are waveform models, the observation would not be possible, since the data model alone cannot identify the source.<sup>7</sup>

It is important to recognize that the entire process constitutes two observations: the observation of gravitational waves and the observation of a binary black hole merger, which emits gravitational waves. With this in mind, the details about detector validation and data processing are not directly relevant to the central argument of the current paper, as it is primarily concerned with the latter observation. However, it is noteworthy that the original data produced by the LIGO detectors, termed “strain data,” is vast and unprocessed. It is subject to various instrumental noises and environmental disturbances, necessitating a complex reduction and cleaning process. In essence, the unprocessed strain data hinders the identification of gravitational wave signals and the subsequent inference of the source system.

Two search techniques are employed in the detection of gravitational wave signals from the data. One is a model-based search, known as matched filtering. It involves the search for matching signals by comparing potential, unspecified signals in the data with pre-defined templates derived from models of the coalescence of binary compact objects based on general relativity. In essence, there is a library of pre-defined gravitational wave templates and an unknown signal. The templates are then used to identify their matches in the unknown signals, a process that primarily aids in determining the amplitude of the detected gravitational wave compared to the noise. It is evident that this observation is theory-laden, as the sensitivity of the detector and the models of the source are taken into account. Scientists anticipate a specific range of signals to be detected and search for them in the unknown detected signal. The other is an unmodeled generic search to identify specific coincident signals within a preferred time and frequency range. That is, combining and comparing data simultaneously detected by different detectors (LIGO Hanford, LIGO Livingston) to identify gravitational wave bursts (short duration signals) purely based on data patterns, as opposed to template-based searches. This also helps to identify the gravitational wave signal relative to noise but is not sensitive enough to identify the source system. Both searches are meant to identify candidate events that their likelihood is high enough to be considered as gravitational wave signal. That is, events whose detection-statistic value is above that of the detector noise. The detection statistic is a numerical measure used to distinguish real signal from noise. It is a threshold above which any signal that exceeds it is considered unreliable. This is a complicated and challenging process, as it is rather impossible to determine the exact value of background noise. However, the main factors considered to determine this value are background noise estimation, injections (false signals), and statistical significance: the probability of an event occurring by chance.

The results of these two different searches are consistent with each other, and this is an indication that the origin of the signal is astrophysical. Thus, the collaboration report that “waveform analysis of this event indicates that if it is astrophysical in origin, it is also a binary black hole merger” (Abbott et al., 2016). Now the question is, how do we know that it is astrophysical in origin and it is a binary black hole merger?

---

<sup>7</sup>For details about the modeling process and their limitations see (Pretorius, 2005; Schmidt, 2020; Blanchet, 2014).

To answer this question, one needs to unearth the details of the epistemic aspect of this observational process. There is no clear distinction, but for the sake of clarity one can break down the question into two parts: one is about the validity and reliability of the detector and the information channel, and the other is about the justifiability of the inferences employed in discerning the source of the information. Here, I am concerned with the second part assuming that the event is astrophysical.

Yunes and Pretorius (2009) distinguish two sources of error: modeling bias and fundamental bias. The former concerns the inaccuracy of the models with respect to GR, and the latter concerns the inaccuracy of GR with respect to the source system. Elder argues in detail that the model-ladenness of the binary black hole merger observation contains a circularity, namely, “the accuracy of the models must be established using the LIGO-Virgo observations, but these observations *assume* the accuracy of the models” (2023, p. 3, emphasis in original).<sup>8</sup> Furthermore, the most pressing issues regarding binary black hole mergers, she argues, are the following. First, the binary compact objects are inaccessible in terms of manipulation and intervention. That is, we cannot calibrate the interferometry against them. Second, there is no alternative channel to access them. Third, the phenomenon represents a regime that has not yet been tested, so the previous success of GR does not guarantee that it will be correct in this regime as well.<sup>9</sup> She further argues that this is crucial because inaccuracies in the models can lead to inaccuracies in the observations, and thus to a mischaracterization of the system. Justifying the accuracy of the models is a prerequisite for unbiased observation.

One should note that this problem goes beyond the experimenter’s regress introduced by Collins (1981), i.e., even if we break the circularity with respect to the instrument, as Franklin (1994) argues for, there is still a concern about observing binary black holes. And most importantly, Franklin’s treatment cannot explain it.

Cutler and Vallisneri (2007) raise epistemic uncertainties regarding the possibility of model inaccuracy in the observation of binary compact mergers. This is due to the assumptions and simplifications behind and in the model construction. For example, the assumption that all binaries have circularized before merging, or unverified assumptions about the accuracy of the model itself. In particular, they focus on inaccuracies in post-Newtonian approximations as models of the inspiral phase.

Patton argues that if the model assumes that a phenomenon has a certain property, there is no guarantee that deviations from this assumption can be detected during testing. However, she acknowledges that the significance of this problem is contingent, i.e., relative to the theory in question (2020, p. 143).

Yunes and Pretorius state that there is plenty of observational evidence that extremely compact objects exist as predicted by general relativity, for example in the center of our Milky Way, so the crucial question is not whether black holes exist or not, but whether GR can accurately describe them. However, they raise two very important questions, namely,

---

<sup>8</sup>The alleged circularity in the observation of binary black holes bears a striking resemblance to the alleged circularity in Galileo’s observation of the Moon’s mountains. For discussions of the latter, see (Feyerabend, 1988; Ariew, 1984; Shea, 2000; Spranzi, 2004).

<sup>9</sup>As I will show later, the “guarantee” and “correctness” conditions are not necessary.

(i) Suppose gravity is described by a theory differing from GR in the dynamical, strong-field regime, but one observes a population of merger events filtered through a GR template bank. What kinds of systematic errors and incorrect conclusions might be drawn about the nature of the compact object population due to this fundamental bias? (ii) Given a set of observations of merger events obtained with a GR template bank, can one quantify or constrain the level of consistency of these observations with GR as the underlying theory describing these events? (2009, p.3).

Since GR is not the only candidate theory of gravity, but e.g. Brans-Dicke theory, massive graviton theory, Chern-Simons modified gravity, Einstein ether theory, MOND, TeVes, and DGP theory are proposed as alternatives to it (Hess, 2020), Yunes and Pretorius (2009) and Baker et al. (2017) present results on excluded alternatives based on LIGO observations.<sup>10</sup>

### 3.1 Parameterized post-Einsteinian Framework

An effective strategy which is normally being employed to justifiably delimit the possible theoretical models, is generalizing and expanding the scope of the fundamental assumptions to the extent which go beyond the specific general background theory on which the theoretical models are built. Massimi calls it “theoretical constraints” or “pre-existing, independent theory with a broader scope of application” (2007, pp. 250 - 251).<sup>11</sup> This approach allows the eliminative reasoning to justifiably rule out the number of alternatives on a scale which include alternatives that are based on other theories. The idea is also in line with what Norton (1995) calls assumptions and arguments that are general enough and uncontroversial enough to encompass a large number of alternatives. Interestingly, this kind of strategy is being used to mitigate the concern regarding deviations from general relativity in the observation of binary black holes, and is called parameterized post-Einsteinian.<sup>12</sup>

As we have seen, the main concern regarding the observation of binary black holes is as follows. General relativity is an untested theory in the extreme dynamical regime where the merger of compact binaries takes place. However, the modeling of these compact binaries assumes that Einstein’s field equations are correct. This could lead to an incorrect inference of the exact binary compact based on these models, because if GR does not adequately describe these binaries in this regime, their waveform models could deviate from GR. Since model-based searches are crucial to binary black hole observations, the set of models/templates does not contain any possible event that

---

<sup>10</sup>Note that the results presented by Baker et al. (2017) are based on multi-messenger detection of gravitational waves from binary neutron stars. It could be argued that this type of observation does not rely primarily on model assumptions.

<sup>11</sup>The best explanation for Massimi’s use of the term “independent” has to do with a high-level theory that cannot be directly tested against the data, not the independent argument promoted by Hacking, Kosso, and the others. Otherwise, the theory is by definition involved in the observational process through its models.

<sup>12</sup>I am grateful to reviewer one for drawing my attention to a similar parameterization technique for the dark energy equation of state, see Wolf and Ferreira (2023) for details about its limitations. Another approach that has recently become popular in cosmology to explain the expansion of the universe is called Horndeski gravity. However, unlike the parameterization approaches used in gravitational waves and dark energy cases, it is a generalized framework or family of modified gravity theories that introduces a scalar field while maintaining second order equations of motion. For the original formulation of the theory by Gregory Horndeski see Horndeski (1974), and for a recent review see Horndeski and Silvestri (2024).

deviates from GR. This raises the concern that there may be systematic errors in the conclusions drawn about the source of the detected gravitational waves.

To address this problem, Yunes and Pretorius develop a strategy called parameterized post-Einsteinian (ppE). Although the framework is proposed to be used to identify deviations after the detection process based on GR templates has been carried out (2009, p. 5), it is worth considering, since it was proposed years before the observation of binary black holes and it is a strategy to alleviate the concerns discussed in the bulk of this paper. It is a formalized framework for recognizing the possibility of dependencies and independencies of parts and assumptions of a theory upon each other, in order to provide a justifiable test of an intended part of that theory.

Yunes and Pretorius find it reasonable to start with binary black hole mergers, an event within the scope of GR, based on the evidence they provide for the remarkable success of GR in many tests that have been thrown at it from its conception to the present. For example, a large amount of observational evidence in favor of binary compacts as described by GR, black holes at the center of galaxies, X-ray binary systems, and binary pulsars (2009, p. 2).<sup>13</sup> They try to strike a balance between modifying GR to the extent that possible deviations close to the theory could be identified, and at the same time they keep the modifications minimal enough to be able to analyze the data. For example, they identify a set of physical systems in the merging phase by identifying those claims and assumptions of GR that are generic enough to be applied to this set of systems as constraints on them in this phase.<sup>14</sup>

The parameterization approach is a counterfactual strategy in which the variables and assumptions of the model are parameterized to track their sustainable range. This results in the generation of counterfactual waveforms (theoretical models) whose predictions may differ from GR.<sup>15</sup> For example, the assumption of “continuity” of the waveform through the merging and ring-down phases is explored by introducing a set of parameterized constants that are governed by this assumption. This technique is also used to parameterize the governing equations and dynamics of the black hole merging processes to identify stable and varying formal parts of the theory. As noted in the quote above, all of these modifications and explorations to expand the scope of testing theories of gravity in this dynamical strong-field regime are based on more general and relaxed assumptions.

### 3.2 Limitations of the Parameterized Post-Einsteinian Framework

Again, recall what Norton (1995) calls assumptions and arguments that are general enough and uncontroversial enough to encompass a large number of alternatives, and what Massimi (2007) calls “theoretical constraints” or “pre-existing, independent theory with a broader scope of application.” This demonstrates how similar their mode of reasoning is to that presented by Yunes and Pretorius (2009).<sup>16</sup> That is, it appears that

---

<sup>13</sup>Note that this consideration by Yunes and Pretorius supports my claim in the following sections about the integration of meta-empirical assessment into bidirectional inference to the source.

<sup>14</sup>See Perkins and Yunes (2022) for more recent analyses of the robustness of the framework.

<sup>15</sup>One may also find similarities to counterfactual causal reasoning developed in Woodward (2003).

<sup>16</sup>Patton (2020) argues that in terms of theory testing and theory confirmation, the framework is similar to methodological approaches to theory testing in philosophy of science such as those of Rudolf Carnap,



the ppE framework is a strategy developed to provide a broader space of alternatives in the first step of eliminative reasoning, which in turn provides a more robust basis for eliminating alternative theories of gravity in the dynamical strong-field regime based on the empirical evidence gathered through the detection of gravitational waves in the second step.

However, as they point out, the ppE framework is developed only to identify deviations from GR for binary black hole mergers and the viable proposed binary compacts such as boson stars that emit similar waves as binary black holes (2009, p. 6). Furthermore, as they also point out, at the level of theoretical models (templates), the ppE framework cannot provide an exhaustive number of possible models, nor is it a framework capable of identifying an exhaustive list of deviations from GR to exclude all the possible alternatives to GR by providing its modifications based on more general principles and assumptions of the theory itself (2009, pp. 2-4).

Let me reiterate the main points about the limitations of ppE. First, it is specifically designed for post-detection analyses, and mostly for testing GR in a more robust way, rather than for reliably inferring the binary black hole merger. Furthermore, for post-detection analyses to work, the ppE modifications are kept minimal to be able to analyze the detected data. Thus, they argue,

We do not propose here to employ the ppE templates for direct detection, but rather for post-detection analysis. Following the detection of a GW by a pure GR template, that segment of data could then be reanalyzed with the ppE templates (2009, p. 5).

Second, GR assumptions are not just used for determining parameters of the source system, they are also underlying assumption for the bank of templates used in matched filtering to detect gravitational waves. So, any kind of signal that is not modeled and not expected by the bank of templates might systematically be disregarded. Furthermore, as indicated before, it can only identify deviations from GR for binary black holes and those proposed sources of gravitational waves that mimic them. Also, the parameterized models that are built based on this framework are not exhaustive therefore it cannot address the problem of unconceived alternatives. As it appears in the following quotation Yunes and Pretorius admit that one can modify GR in “uncountably many conceivable” ways, not to mention the unconceivable ways and unconceivable theories that do not satisfy the criteria they employ as guidance.

In theory, there are uncountably many conceivable modifications to GR that only manifest in the late stages of the merger. To make this question manageable, we shall guide our search for ppE expansions by looking to alternative theories that satisfy as many of the following criteria as possible: (i) *Metric theories of gravity*: theories where gravity is a manifestation of curved spacetime, described via a metric tensor, and which satisfies the weak equivalence principle. (ii) *Weak-field consistency*: theories that reduce to GR sufficiently when gravitational fields are weak and velocities are small, i.e. to pass all precision, experimental, and observational tests. (iii) *Strong-field inconsistency*: theories that modify GR in the dynamical strong field by a sufficient amount to observably affect binary merger waveforms (2009, p. 2, emphasis in original).

---

Carl Gustav Hempel, and Howard Stein. Similar ideas can be found in Glymour (1975) on theory testing through the method of bootstrapping.

Considering that, one needs to provide further reasons why one is justified in not exploring the full space of possible binary compact objects, or any other unknown entity that might mimic black holes in emitting gravitational waves, and why one is justified in not exploring the full space of gravitational theories which might in turn represent different sources of gravitational waves. As Dawid points out, the canonical view of the scientific method, based on its narrow view of evidence, cannot explain why, in practice, scientists increase their confidence in theories to be successful in future tests when they have a successful record in the past (2018, p. 496). Similarly, the immediate acceptance of the observation of binary black holes by the scientific community, despite the epistemological concerns mentioned above, requires an explanation. That is, within the canonical understanding of the concept of observation, the source of gravitational waves is inherently underdetermined by the information we gather from the empirical evidence (the waves) and the information we gather from the theoretical constraints (e.g., ppE).

## 4 Back to IOE: Towards a Reliable Inference to the Source

It may now be clear that these concerns of scientists, philosophically reconstructed by Elder (2023) as concerning circularity in the case of the observation of binary black holes,<sup>17</sup> fall into the same categories of concerns raised by Hacking, Kosso, Azzouni and Franklin, namely the involvement of the source's theory/model in the observational process. Thus, to conceive of the situation as circularity is to interpret the observational process within the framework of hypothetico-deductivism. In what follows, I defend the claim that the version of the inference to the only explanation advocated by Massimi (2007) captures the situation well and mitigates the concern about the inaccuracy of the models with respect to GR. However, to fully establish an epistemic justification for the observation regarding the viability of GR, or even parameterized post-Einsteinian frameworks, as a background theoretical framework and its accurate description of the source system, one needs to complement Massimi's with a broader conception of evidence. To this end, I will argue that the introduction of meta-empirical assessments, as in McCoy (2021), to support eliminative reasoning in the observational context is a promising strategy as a justifiable extension of the concept of evidence.

### 4.1 Bidirectional Inference to the Only Explanation: The Case of Binary Black hole Merger

In this subsection, I begin by drawing similarities between Massimi (2007) and the binary black hole inference. The reasoning could account for the inference if one considers only those possible alternative sources that are located within GR or, more precisely, within parameterized post-Einsteinian frameworks.

---

<sup>17</sup>For a description of different senses of circularity and the limits of their epistemic concerns, see (Evans and Thébaud, 2020).

Patton distinguishes between two sets of parameters in the LIGO methodology, the first set being the “hypothetical physical parameters” used to construct numerical relativity simulations, and the second set being the “estimated physical parameters” inferred from the detected waveform. The search process of finding candidate waveforms in the data is laden with theoretical assumptions, i.e. the first set. The hypothetical dynamic characteristics of the systems, such as the “chirp mass” as a parameter, are built by considering a range of hypothetical physical values. That is, a set of theoretical models of the waveform is built. These hypothetical waveforms are used in the search to find the most likely fitting signal. To track the change in the physical parameters of the system, the chirp mass plays an important role because it is time-covariant with those parameters (2020, p. 148).

Based on that, it appears that the process of choosing candidate waveforms which best fits is from both sides, namely bidirectional. In a simpler way, there is a raw data (strain data) which plays the role of a data model providing a range of possible expected values based on its sensitivity, there are theoretical waveform models, then a signal is extracted using matched filtering and burst pipelines. Finally we relate the properties of the signal (like frequency, polarization, and strength) to the properties of the source (like mass, spin, chirp mass, orbital frequency, distance, and position).<sup>18</sup> The source manifests itself from the data model resulting from a reliable detector and from the theoretical models constrained by theoretical background and principles; eventually there is a convergence towards selecting the most appropriate candidate. The following quotations from the discovery paper illustrate the reasoning:

The most plausible explanation for this evolution is the inspiral of two orbiting masses,  $m_1$  and  $m_2$ , due to gravitational-wave emission (2016, p 3).

Also,

To reach an orbital frequency of 75 Hz (half the gravitational-wave frequency) the objects must have been very close and very compact; equal Newtonian point masses orbiting at this frequency would be only  $\approx 350$  km apart. A pair of neutron stars, while compact, would not have the required mass, while a black hole neutron star binary with the deduced chirp mass would have a very large total mass, and would thus merge at much lower frequency. *This leaves black holes as the only known objects compact enough to reach an orbital frequency of 75 Hz without contact*(2016, p. 3, emphasis added).

The above inference process is very similar to what Massimi (2007) reconstructed in the case of the observation of the  $J/\psi$  particle, which she called “saving unobservable phenomena.” That is, in an upward path, from the data appears an evolution represented by the frequency of the signal, and this indicates an inspiral-merging-ringdown process of two compact bodies. In a downward path, theoretical entities (binary compacts) introduced by theoretical model(s) are eliminated, and finally all of them are eliminated, saving one.

One should not conflate this bidirectional inference to the only explanation with what Hasok Chang calls “epistemic iteration.” Epistemic iteration is not bidirectional inference in the sense of inferring the source of detected empirical data from both

---

<sup>18</sup>Relating the properties is carried out by using Bayesian formalism.

models of the data (upward) and models of the source (downward) in a single observational setting or for the purpose of observing a particular phenomenon. Epistemic iteration, in Chang’s understanding, is rather a successive process in which scientists revisit the knowledge claims in science to arrive at an ever-improving result. It is usually not tied to a single measurement or observational process but continues over time. For example, Chang shows that confidence in the reliability of thermoscopes was built by comparing this instrumental result with human sensations of temperature, but later it turned out that precise thermoscope measurements could correct sensory perceptions. This is a progressive process, taking into account standards such as accuracy, consistency, simplicity, explanatory power, and so on (Chang, 2004).<sup>19</sup>

However, as indicated before, there is a legitimate concern that Massimi does not treat properly in her paper, and which reappears in our example. Admittedly, this bidirectional inference process may legitimately pick up a theoretical entity as the only explanation, but without a well-founded and justified constraint on the number of theoretical models and hence entities, there is a concern that the selected explanation in terms of a theoretical entity may not be the correct one. The correct model that introduces the correct entity might locate outside the considered set. In this case, the selection process will be successful, but the selected entity will be an incorrect representation of the evidence. For example, the models she discussed in her paper, including the Gell-Mann-Zweig quark model which saved the phenomenon, are based on the background quark theory, which in turn follows some principles and theoretical assumptions. If these principles and assumptions are not justified, then the derivation based on them is also not justified.

Or in our case study, the derived value of the chirp mass, on the basis of which other binary compacts are eliminated, is based on theoretical waveforms that in turn are based on the linearized Einstein field equations in the strong field regime, NR simulations, and the quadrupole formula (Capano et al., 2016; Yunes and Pretorius, 2009). It is necessary to provide a justification for these GR-based theoretical assumptions that have constrained the model waveforms.

## 4.2 Meta-Empirical Support for the Inference to the Source

Recall that van Fraassen (1985, 2001, 2002) extends the realm of observables to images produced by instruments, but remains agnostic about the source representing them, since there is no sensory access to the source. The only plausible explanation for this rejection is that the source is not demonstrable, otherwise it is both inconsistent to extend the realm of observables to instrumental images based on scientific results without fully endorsing the scientist’s use of the concept of observation, and it is a misunderstanding of the concept of phenomenon (pace Bogen and Woodward) if he restricts the concept of phenomenon to instrumental images. Thus, I argue, the central concern in van Fraassen’s rejection of extending the concept of observation to the source of instrumental images stems from the underdetermination thesis. Supporting evidence for this is his “best of bad lot” objection to IBE, which can be briefly expressed as follows. Assuming that selecting the best explanation from the set of

---

<sup>19</sup>I am grateful to reviewer two for bringing this point to my attention.

available explanations is justified, the inference is not justified because it is generally likely that the true explanation is a member of the unconceived alternatives (van Fraassen, 1989). Since IBE and Eliminative Reasoning are both two-step reasoning and differ only in the second step, and van Fraassen’s objection is directed at the first step of IBE, it can be extended to Eliminative Reasoning as well.

The version of eliminative reasoning that I have so far advocated in the case of the inference to the source in the process of observation is a bidirectional one. That is, indication about the source can be traced from the empirical evidence represented as data models (bottom-up) and from the theoretical models derived from a background theory (top-down). I have argued that in order for the bidirectional eliminative reasoning to work for the inference to the source in the second step, one needs to justify delimiting the number of alternatives in the first step. That is, given the reliability of the detector and the process of constructing the data model, an epistemic issue which I have tried to explore throughout the article is the justifiability of the basis for taking the background theory and its possible extensions as theoretical constraints (pace Norton and Massimi) to delimit the number of alternatives one can consider for the source. In this section, I argue that without extending the concept of “evidence” beyond the empirical to give epistemic reasons, delimiting the alternatives based on background theory or its extensions as theoretical constraints is epistemically unfounded.

Stanford (2006) identifies this problem in eliminative reasoning generally and McCoy (2021) suggests to resolve it by invoking meta-empirical assessments developed by Dawid (2013, 2016, 2018).

Dawid’s methodology of meta-empirical theory assessment was originally intended to address the situation in theoretical fundamental physics where, despite the lack of direct empirical evidence, scientists have built up a high degree of confidence in these theories. The exemplary case is string theory, which is proposed to provide a universal description for all known physical interactions, including gravity, by introducing the idea of strings replacing point particles. The methodology of meta-empirical assessment challenges canonical theory confirmation, which focuses primarily on direct empirical evidence, by identifying three meta-empirical assessments that put particularly strong constraints on scientific underdetermination (assessing how constrained theory space is). In other words, where scientific underdetermination is severely constrained, the likelihood that a theory is empirically adequate increases (Dawid, 2013). To this end, Dawid (2013) has identified three meta-empirical observations that support the statement that underdetermination is strongly limited: First, the No Alternatives Argument (NAA), which is the observation that scientists have not found alternatives to the current theory despite their extensive search. Second, the Meta-Inductive Argument (MIA), is the observation that theories in the research program that meet some general set of criteria have a predictively successful history. Third, the Unexpected Explanatory Argument (UEA), is the observation that the explanatory scope of the theory expands considerably beyond the scope for which it was originally developed. It is crucial to note that meta-empirical theory assessment preserves the role of empirical evidence that can support a theory within its empirical domain, where available. The meta-observations are a type of evidence that is not predicted by the theory, but nevertheless provides evidence for the viability of the theory. The

significance of meta-empirical assessments has been identified in various episodes and parts of science as an implicitly integrated part to the methodology of theory assessment (Dardashti et al., 2019; Dawid, 2021; McCoy, 2021; Baerdemaeker and Dawid, 2022; Dawid and McCoy, 2023; Wolf, 2024). The present paper also contributes to this literature by identifying their significance in the observation of a binary black hole merger.

It is now clear that the bidirectional eliminative reasoning active in the process of observation is a species of general eliminative reasoning and its only specific feature is a bidirectional inference to the source in the second step. Therefore, it is reasonable to extend the treatment suggested by McCoy to justify the inference to the source in the observation of binary black holes via gravitational waves.

He poses the problem in the form of a dilemma. That is, on the one hand, if one argues for an epistemic justification for eliminative reasoning, one should epistemically justify the premises of greater generality (delimiting the alternatives); on the other hand, if one accepts that delimiting the alternatives is done pragmatically, one should bite the bullet and accept that the conclusion of eliminative reasoning is not epistemically justified because it depends on pragmatic assumptions. He argues that proponents of eliminative reasoning, including John Norton and John Dorling, have failed to give epistemic reasons for their claims about calling premises of greater generality “beyond reasonable doubt” and “plausible” respectively (2021, pp. 7-8). By extension, in the case of binary black hole observations, the dilemma continues: if one infers binary black holes on the basis of narrowing down the alternatives by assuming that GR and its ppE extension are correct, one must epistemically justify this assumption. Otherwise, if this assumption is pragmatically motivated, the inferential step toward selecting binary black holes as the source of the detected gravitational waves is epistemically unfounded.

Since the general view tends to reject the second horn of the dilemma, the only possible option is to accept the extension of the concept of evidence beyond the experimental result (going beyond the narrow empiricist view of the theory-evidence-confirmation relation in the hypothetico-deductivist conception) in order to overcome the underdetermination of the source by information. Information is to be understood as the evidence we gather from both empirical (data model) and theoretical (theoretical model) vessels in the bidirectional process of the inference to the source in the second step of eliminative reasoning. That is, we must provide epistemic reasons for believing that there are no other valid alternatives beyond the considered set of alternatives in the context of the inference to the binary black holes.

Why the assumption that it is completely implausible that the source of gravitational waves is something outside of (deviating from) the basic assumptions of GR, even though the theory has not been tested in the extreme gravitational regime which is assumed to be the source of the waves? Recall the direct statement I quoted above from the Discovery paper, viz: “This leaves black holes as *the only known objects* compact enough to reach an orbital frequency of 75 Hz without contact” (2016, p. 3, emphasis added). The assumption is based on the assessment of the limitations to scientific underdetermination. In other words, how many empirically distinguishable alternatives can explain the data (the gravitational waves)? What kind of evidence

can be provided to support this assumption? The most plausible answer, I argue, is that there must be some evidence for GR that goes beyond the empirical domain of the theory, i.e., meta-empirical evidence that supporting the hypothesis that there are no alternatives which can deviate from the basic assumptions of GR in the intended domain (the extreme gravitational regime). There seems to be no reasonable doubt that the basic assumptions of GR satisfy all three of Dawid’s meta-empirical evidence. As noted above, the research program has been remarkably successful in surviving various crucial tests (an incredibly successful history that can be thought of as MIA).<sup>20</sup> A century after the development of the theory, there is no viable alternative theory for the intended domain of application that violates the basic assumptions of relativity and has passed all the tests (NAA).<sup>21</sup> The theory is known for making a considerable number of novel predictions which can be conceived as Unexpected Explanatory Arguments (UEA). These observations provide epistemic evidence for the hypothesis that there are no alternatives which can deviate from the basic assumptions of GR in the intended domain (the extreme gravitational regime), thus stating strong limitations to scientific underdetermination, otherwise the success of the general relativistic paradigm remains unexplained.

One might argue that ppE encompasses a wide variety of modified gravity models in the vicinity of GR, i.e. those models that are metric theories and are considered viable options to pursue. Thus, the concern about gravity theories that do not assume a metric framework as their basis, let alone unconceived alternatives, is not justified since we have a substantial amount of evidence for GR. Therefore, any modified gravity theory must remain close to GR to account for its remarkable success.

The objection implicitly invokes Dawid’s meta-inductive argument, and from it argues for strong limitations on scientific underdetermination. That is, the meta-observation that GR has been remarkably successful is used to confirm the claim that any future theory of gravity should account for the successful history of GR, which in turn imposes strong limitations on scientific underdetermination. Otherwise, the empirical evidence for predictions derived from GR in its domain of application cannot support the claim that there are no alternatives to GR, so the objection fails.

Moreover, in the case of the observation of binary black holes, a probabilistic significance of the epistemic significance of this meta-empirical evidence is sufficient, because the full epistemic justification of the observational claim does not come only from the top-down approach, but is a corroborative process from both directions. Therefore, a probabilistic epistemic significance to justify the claim about the limitations to scientific underdetermination (establishing a high probability for the viability of the theory) can do the job for the first step of eliminative reasoning.

---

<sup>20</sup>Wolf (2024) uses gravitational waves as an example of meta-empirical reasoning (MIA argument), demonstrating how the exceptional success of the general relativity (GR) research program and its models justified physicists’ conviction that gravitational waves existed long before they were empirically detected. I am grateful to reviewer one for drawing my attention to this paper.

<sup>21</sup>One could disagree on this point, see Wolf et al. (2024) for a comprehensive review of underdetermination with respect to classical and modern tests of GR.

## 5 Conclusion

The philosophy of science literature on observation is divided into two camps, the empiricist camp insisting on restricting the concept of observation to perception, and the practice-oriented liberal camp arguing for extending the concept of observation to include theoretical entities represented by instrumental results.

I believe that the empiricist camp is untenable for several reasons. For example, although the waveform is only a representation of the coalescence of a binary black hole, according to the most sophisticated empiricism, constructive empiricism, one is forced to consider the instrumental result of LIGO and Virgo, the waveform, as the phenomenon to be saved, and at the same time one is forced to consider the binary as unobservable. Paradoxically, the binary would only be observable if a human agent had been able to perceive it. Moreover, constructive empiricism is alien to the scientist's use of the concept of observation, so it is an uninteresting position for someone trying to understand the concept of observation in practice.

At the same time, the practice-oriented liberal camp has failed to provide an account of observation that both accounts for as many cases in practice as possible and preserves its epistemological ground. The inference to the source of gravitational waves is a case in point. The liberal accounts of observation are either so narrow that they fail to account for it, or so loose that the epistemic aspect of this observation becomes questionable.

The present paper has attempted to unfold the above difficulties and to identify flaws in the methodological frameworks used to reconstruct the “inference to the source” step in observational processes by available philosophical accounts of observation. It has also been argued that the most plausible way is to embed this inferential step in an eliminative framework, but this necessarily requires an extension of the concept of evidence.

## Declarations

**Acknowledgments** I would like to thank the members of the “Epistemology of Modern Physics” online group, Jeremias Düring and Muhsin Burhan, for their valuable comments and discussions on earlier drafts of this paper. Thanks to the participants of the “Establishing the Nordic Network for Philosophy of Physics Workshop,” the “Methodological Transformations in Fundamental Physics Workshop,” the “Philosophy of Science Seminar at the University of Bristol,” and the “Philosophy of Physics Seminar at the University of Oxford” for their comments on presentations based on this paper. I also benefited from discussions with Sam Fletcher and Karim Thébault on the central arguments of the paper, and from Lydia Patton, who provided written comments on the slides I used in my presentation. Special thanks to Radin Dardashti for discussions and detailed comments on earlier drafts, and two anonymous reviewers for valuable comments.

**Funding** This research is funded by the DFG Research Training Group 2696.

**Conflict of interest** The author confirms that all conceptualization, analysis, and writing of this article were conducted independently. No financial, institutional, or personal relationships influenced the content of this work.



## References

- Abbott, B.P. et al. 2016, Feb. Observation of gravitational waves from a binary black hole merger. *Physical Review Letters* 116(6): 061102. <https://doi.org/10.1103/PhysRevLett.116.061102> .
- Adam, M. 2004. Why worry about theory-dependence? circularity, minimal empiricity and reliability. *International Studies in the Philosophy of Science* 18(2 & 3): 117–132. <https://doi.org/10.1080/0269859042000296486> .
- Ariew, R. 1984. Galileo’s lunar observations in the context of medieval lunar theory. *Studies in History and Philosophy of Science Part A* 15(3): 213–226. [https://doi.org/10.1016/0039-3681\(84\)90017-7](https://doi.org/10.1016/0039-3681(84)90017-7) .
- Azzouni, J. 1997. Thick epistemic access. *Journal of Philosophy* 94(9): 472–484. <https://doi.org/10.5840/jphil199794926> .
- Azzouni, J. 2004. Theory, observation and scientific realism. *British Journal for the Philosophy of Science* 55(3): 371–392. <https://doi.org/10.1093/bjps/55.3.371> .
- Baerdemaeker, S.D. and R. Dawid. 2022. Mond and meta-empirical theory assessment. *Synthese* 200(5): 1–28. <https://doi.org/10.1007/s11229-022-03830-8> .
- Baker, T., E. Bellini, P.G. Ferreira, M. Lagos, J. Noller, and I. Sawicki. 2017, Dec. Strong constraints on cosmological gravity from gw170817 and grb 170817a. *Phys. Rev. Lett.* 119: 251301. <https://doi.org/10.1103/PhysRevLett.119.251301> .
- Bird, A. 2005. Abductive knowledge and holmesian inference, In *Oxford Studies in Epistemology*, eds. Gendler, T.S. and J. Hawthorne, 1–31. Oxford University Press.
- Bird, A. 2007. Inference to the only explanation. *Philosophy and Phenomenological Research* 74(2): 424–432. <https://doi.org/10.1111/j.1933-1592.2007.00028.x> .
- Bird, A. 2010. Eliminative abduction: Examples from medicine. *Studies in History and Philosophy of Science Part A* 41(4): 345–352. <https://doi.org/10.1016/j.shpsa.2010.10.009> .
- Blanchet, L. 2014. Gravitational radiation from post-newtonian sources and inspiralling compact binaries. *Living Reviews in Relativity* 17: 2. <https://doi.org/10.12942/lrr-2014-2> .
- Bogen, J. and J. Woodward. 1988. Saving the phenomena. *Philosophical Review* 97(3): 303–352. <https://doi.org/10.2307/2185445> .
- Brown, H.I. 1993. A theory-laden observation can test the theory. *The British Journal for the Philosophy of Science* 44(3): 555–559 .

- Brown, H.I. 1994. Circular justifications. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1994: 406–414 .
- Cabrera, F. 2017. Can there be a bayesian explanationism? on the prospects of a productive partnership. *Synthese* 194(4): 1245–1272. <https://doi.org/10.1007/s11229-015-0990-z> .
- Capano, C., I. Harry, S. Privitera, and A. Buonanno. 2016, Jun. Implementing a search for gravitational waves from binary black holes with nonprecessing spin. *Phys. Rev. D* 93: 124007. <https://doi.org/10.1103/PhysRevD.93.124007> .
- Carrier, M. 1989. Circles without circularity, testing theories by theory-laden observations in an intimate relation. *studies in the history and philosophy of science. Boston Studies in the Philosophy of Science* 116: 405–428 .
- Chalmers, A. 2003. The theory-dependence of the use of instruments in science. *Philosophy of Science* 70(3): 493–509. <https://doi.org/10.1086/376924> .
- Chang, H. 2004. *Inventing Temperature: Measurement and Scientific Progress*. New York, US: OUP Usa.
- Collins, H.M. 1981. Son of seven sexes: The social destruction of a physical phenomenon. *Social Studies of Science* 11(1): 33–62. <https://doi.org/10.1177/030631278101100102> .
- Cutler, C. and M. Vallisneri. 2007, Nov. Lisa detections of massive black hole inspirals: Parameter extraction errors due to inaccurate template waveforms. *Phys. Rev. D* 76: 104018. <https://doi.org/10.1103/PhysRevD.76.104018> .
- Dardashti, R., R. Dawid, and K. Thébault. 2019. *Why Trust a Theory?: Epistemology of Fundamental Physics*. New York, NY: Cambridge University Press.
- Dawid, R. 2013. *String Theory and the Scientific Method*. Cambridge University Press.
- Dawid, R. 2016. Modelling non-empirical confirmation, In *Models and Inferences in Science*, eds. Ippoliti, E., F. Sterpetti, and T. Nickles, 191–205. Cham: Springer. [https://doi.org/10.1007/978-3-319-28163-6\\_11](https://doi.org/10.1007/978-3-319-28163-6_11).
- Dawid, R. 2018. Delimiting the unconceived. *Foundations of Physics* 48(5): 492–506. <https://doi.org/10.1007/s10701-017-0132-1> .
- Dawid, R. 2021. The role of meta-empirical theory confirmation in the acceptance of atomism. *Studies in History and Philosophy of Science Part A* 90: 50–60 .
- Dawid, R. and C. McCoy. 2023. Testability and viability: Is inflationary cosmology ?scientific?? *European Journal for Philosophy of Science* 13(4): 51. <https://doi.org/10.1007/s13194-023-00556-3> .

- Dorling, J. 1973. Demonstrative induction: Its significant role in the history of physics. *Philosophy of Science* 40(3): 360–372. <https://doi.org/10.1086/288537> .
- Elder, J. 2023. Black hole coalescence: Observation and model validation, In *Working Toward Solutions in Fluid Dynamics and Astrophysics: What the Equations Donot Say*, eds. Patton, L. and E. Curiel, 79–104. Springer Verlag.
- Evans, P.W. and K.P.Y. Thébault. 2020. On the limits of experimental knowledge. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 378(2164): 20190235. <https://doi.org/10.1098/rsta.2019.0235> .
- Feyerabend, P. 1988. *Against Method*. London: New Left Books.
- Forber, P. 2011. Reconceiving eliminative inference. *Philosophy of Science* 78: 185–208. <https://doi.org/10.1086/659232> .
- Franklin, A. 1994. How to avoid the experimenters’ regress. *Studies in History and Philosophy of Science Part A* 25(3): 463–491. [https://doi.org/10.1016/0039-3681\(94\)90062-0](https://doi.org/10.1016/0039-3681(94)90062-0) .
- Franklin, A. 2002. *Selectivity and Discord: Two Problems of Experiment*. University of Pittsburgh Press.
- Glymour, C. 1975. Relevant evidence. *Journal of Philosophy* 72(14): 403–426. <https://doi.org/10.2307/2025011> .
- Hacking, I. 1983. *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. New York: Cambridge University Press.
- Hess, P.O. 2020, September. Alternatives to Einstein’s General Relativity Theory. *Progress in Particle and Nuclear Physics* 114: 103809. <https://doi.org/10.1016/j.pnnp.2020.103809> .
- Horndeski, G.W. 1974, September. Second-order scalar-tensor field equations in a four-dimensional space. *International Journal of Theoretical Physics* 10(6): 363–384. <https://doi.org/10.1007/BF01807638> .
- Horndeski, G.W. and A. Silvestri. 2024. 50 years of horndeski gravity: Past, present and future. *International Journal of Theoretical Physics* 63: 38. <https://doi.org/10.1007/s10773-024-05558-2> .
- Kennefick, D. 2007. *Traveling at the Speed of Thought: Einstein and the Quest for Gravitational Waves*. Princeton University Press.
- Kitcher, P. 1993. *The Advancement of Science: Science Without Legend, Objectivity Without Illusions*. New York: Oxford University Press.

- Kosso, P. 1988. Dimensions of observability. *British Journal for the Philosophy of Science* 39(4): 449–467. <https://doi.org/10.1093/bjps/39.4.449> .
- Massimi, M. 2004. What demonstrative induction can do against the threat of under-determination: Bohr, heisenberg, and pauli on spectroscopic anomalies (1921-24). *Synthese* 140(3): 243–277. <https://doi.org/10.1023/b:synt.0000031319.64615.49> .
- Massimi, M. 2007. Saving unobservable phenomena. *British Journal for the Philosophy of Science* 58(2): 235–262. <https://doi.org/10.1093/bjps/axm013> .
- McCoy, C.D. 2021. Meta-empirical support for eliminative reasoning. *Studies in History and Philosophy of Science Part A* 90: 15–29 .
- Norton, J. 1987. The logical inconsistency of the old quantum theory of black body radiation. *Philosophy of Science* 54(3): 327–350. <https://doi.org/10.1086/289387> .
- Norton, J. 1994. Science and certainty. *Synthese* 99(1): 3–22. <https://doi.org/10.1007/bf01064528> .
- Norton, J. 1995. Eliminative induction as a method of discovery: How einstein discovered general relativity, In *The Creation of Ideas in Physics*, ed. Leplin, J., Volume 55 of *The University of Western Ontario Series in Philosophy of Science*. Dordrecht: Springer. <https://doi.org/10.1007/978-94-011-0037-3>.
- Norton, J. 2000. How we know about electrons, In *After Popper, Kuhn and Feyerabend*, eds. Nola, R. and H. Sankey, 67–97. Dordrecht: Kluwer. [https://doi.org/10.1007/978-94-011-3935-9\\_2](https://doi.org/10.1007/978-94-011-3935-9_2).
- Patton, L. 2020. Expanding theory testing in general relativity: Ligo and parametrized theories. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 69: 142–53 .
- Perkins, S. and N. Yunes. 2022, Jun. Are parametrized tests of general relativity with gravitational waves robust to unknown higher post-newtonian order effects? *Phys. Rev. D* 105: 124047. <https://doi.org/10.1103/PhysRevD.105.124047> .
- Pretorius, F. 2005, Sep. Evolution of binary black-hole spacetimes. *Phys. Rev. Lett.* 95: 121101. <https://doi.org/10.1103/PhysRevLett.95.121101> .
- Salmon, W.C. 2001. Explanation and confirmation: A bayesian critique of inference to the best explanation, In *Explanation*, eds. Hon, G. and S.S. Rakover, Volume 302 of *Synthese Library*. Dordrecht: Springer. [https://doi.org/10.1007/978-94-015-9731-9\\_3](https://doi.org/10.1007/978-94-015-9731-9_3).
- Schmidt, P. 2020. Gravitational Waves From Binary Black Hole Mergers: Modeling and Observations. *Front. Astron. Space Sci.* 7: 28. <https://doi.org/10.3389/fspas.2020.00028> .

- Shapere, D. 1982. The concept of observation in science and philosophy. *Philosophy of Science* 49(4): 485–525. <https://doi.org/10.1086/289075> .
- Shea, W. 2000. Looking at the moon as another earth: Terrestrial analogies and seventeenth-century telescopes, In *Metaphor and Analogy in the Sciences. Origins*, ed. Hallyn, F., Volume 1 of *Origins*. Dordrecht: Springer. [https://doi.org/10.1007/978-94-015-9442-4\\_6](https://doi.org/10.1007/978-94-015-9442-4_6).
- Spranzi, M. 2004. Galileo and the mountains of the moon: Analogical reasoning, models and metaphors in scientific discovery. *Journal of Cognition and Culture* 4(3-4): 451–483. <https://doi.org/10.1163/1568537042484904> .
- Stanford, P.K. 2006. *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. New York: Oxford University Press.
- Suppes, P. 1960. A comparison of the meaning and uses of models in mathematics and the empirical sciences. *Synthese* 12(2-3): 287–301. <https://doi.org/10.1007/bf00485107> .
- Suppes, P. 2009. Models of data, In *Provability, Computability and Reflection*, eds. Nagel, E., P. Suppes, and A. Tarski. Elsevier.
- van Fraassen. 1980. *The Scientific Image*. New York: Oxford University Press.
- van Fraassen. 1985. Empiricism in the philosophy of science, In *Images of Science*, eds. Churchland, P.M. and C.A. Hooker, 245–368. University of Chicago Press.
- van Fraassen. 1989. *Laws and Symmetry*. New York: Oxford University Press.
- van Fraassen. 2001. Constructive empiricism now. *Philosophical Studies* 106(1-2): 151–170. <https://doi.org/10.1023/a:1013126824473> .
- van Fraassen. 2002. *The Empirical Stance*. Yale University Press.
- Wallace, W.A. 1992. *Galileo's logic of discovery and proof: The Background, Content, and Use of His Appropriated Treatises on Aristotle's Posterior Analytics*. Boston, MA, USA: Springer.
- Wolf, W.J. 2024. Cosmological inflation and meta-empirical theory assessment. *Studies in History and Philosophy of Science Part A* 103: 146–158 .
- Wolf, W.J. and P.G. Ferreira. 2023, Nov. Underdetermination of dark energy. *Phys. Rev. D* 108: 103519. <https://doi.org/10.1103/PhysRevD.108.103519> .
- Wolf, W.J., M. Sanchioni, and J. Read. 2024. Underdetermination in classic and modern tests of general relativity. *European Journal for Philosophy of Science* 14(4): 1–41. <https://doi.org/10.1007/s13194-024-00617-1> .

Woodward, J. 2003. *Making Things Happen: A Theory of Causal Explanation*. New York: Oxford University Press.

Woodward, J. 2024. The place of explanation in scientific inquiry: Inference to the best explanation vs inference to the only explanation. <https://philsci-archive.pitt.edu/22998/>. Accessed: 7 Sept. 2024.

Worrall, J. 2000. The scope, limits, and distinctiveness of the method of "deduction from the phenomena": Some lessons from newton's "demonstrations" in optics john worrall. *British Journal for the Philosophy of Science* 51(1): 45–80. <https://doi.org/10.1093/bjps/51.1.45> .

Yunes, N. and F. Pretorius. 2009, Dec. Fundamental theoretical bias in gravitational wave astrophysics and the parametrized post-einsteinian framework. *Phys. Rev. D* 80: 122003. <https://doi.org/10.1103/PhysRevD.80.122003> .