The Epistemological Status of the Direct and Indirect Observation Distinction

Sarwar Ahmed[1*]

[1*]Interdisciplinary Centre for Science and Technology Studies (IZWT), University of Wuppertal.

Corresponding author(s). E-mail(s): sarwarmuhamad84@gmail.com;

Abstract

For various reasons, it has become common wisdom in science that there exists a principled epistemic distinction between direct and indirect observation. In this paper, I present a twofold argument. First, I argue against such a principled epistemic distinction. Second, I highlight a pervasive incongruence between the methodological and epistemological distinctions between direct and indirect observations. My arguments revolve around the idea that it is one thing to make a methodological distinction between observations and another to ascribe epistemic significance to them. I begin by unfolding the historical and philosophical foundations of the distinction, identifying three tenets that have served to sustain the distinction to the present day. I then provide a detailed analysis of two recent philosophical efforts to preserve the epistemic distinction in astrophysics and specific areas of astrophysics, ultimately suggesting that these approaches face significant challenges.

Keywords: observation, direct observation, indirect observation, inference, empiricism, perception

Contents

1 Introduction

Should direct observation always be deemed more reliable and trustworthy in science? Should we always strive for direct observation if it is possible in science? If your intuition is that this should be the case, have you ever questioned the basis of this belief? For various reasons that will be elaborated upon in this paper, it has become common wisdom in science that there exists a principled epistemic distinction between direct and indirect observation. This has led scientists, on many occasions, to attach the qualifier "direct" to certain observations to indicate that it is epistemically relevant rather than a mere reference to the method of the observation. For example, a number of prominent astrophysicists have discussed the existence of black holes in a roundtable in 1998. Almost all of them refer to the terms "direct" and "indirect" in relation to having conclusive evidence for the existence of black holes or not, respectively, without specifying why direct access is considered epistemically superior (Collmar et al., 1998).

In philosophy, not only various empiricists but also practice-oriented philosophers of science who extended the concept of observation to include imperceptible entities have continued to treat the distinction between direct and indirect observation as epistemically significant. I refer in particular to Dudley Shapere's (1982) account of observation, in which he insists that we can observe the center of the Sun *directly* via neutrinos, since the neutrinos interacting with our receptors on Earth have most likely not been interfered with on their way from the center of the Sun. This is contrasted with, according to him, the observation of the center of the Sun via photons, since the photons coming from the center of the Sun encounter numerous interactions. As we shall see in more detail later, on the basis of this contrast, Shapere calls the non-interference construal non-inferential and the other inferential. Shapere's distinction seems appropriate for astrophysical observations. According to his account, the observational process consists of three elements: an observed source, the observer (receptor), and a transmitter of information between the two.

On the same topic, Doboszewski and Lehmkuhl (2023), interpreting Hacking (1989) and Eckart et al. (2017), indicate that because epistemic access to black holes is via indirect observation only, compared to direct observational access to most other astrophysical objects, the existence of black holes might be seen as "somehow less certain and less conclusive" (p. 232).

Perhaps Franklin (2017) was the first to explore this distinction historically. He applies Shapere's distinction to historical cases and concludes that direct observation is not necessarily epistemically superior to indirect observation. Elder (2025) takes a more moderate position by examining the distinction between direct and indirect observations when observing gravitational waves. Like Franklin, she explicitly abandons the ambitious principled epistemic distinction, but at the same time retains it in specific contexts such as gravitational wave observation, with the ambition of broadening its scope. In her epistemic distinction, Elder emphasizes the difference in the nature of justification in the measurement processes of direct and indirect gravitational wave observations. To delineate the structure of justification of each observation, she uses the criterion of "intervention" to determine what constitutes the measuring detector and what constitutes a separate source system. However, as we shall see in detail, Elder appears to oscillate between two criteria - intervention and model-independence - in order to sustain the distinction between direct and indirect observation. The criterion of "intervention" seems to play a primarily methodological role, whereas "model-independence" functions more clearly as an epistemic criterion. This divergence in their respective roles gives rise to a tension within the account.

In this paper, I defend two distinct but related claims, one more general and one more specific, namely: (A) for a principled distinction to be made on epistemological grounds, there is no set of epistemic standards that corresponds to at least a significant number of methodologically (in)direct observations. The epistemic significance is contextually determined, i.e. the (in)directness of the method of observation is not essential for the epistemic assessment. In addition to the reliability of the method, the epistemic evaluation of an observation depends on the reliability of the physical principles, theories and assumptions necessary for the inferences involved. Their reliability is subject to a contextual assessment. (B) When one grounds the distinction on inferential versus non-inferential as an epistemic guide, there is a certain degree of incongruence between determining the inference-free domain and circumscribing the observing instrument or detector.

The paper, therefore, contributes to the literature in two dimensions: First, it extends the line of reasoning initiated by Franklin and Elder by examining the foundations of the principled distinction, tracing its echoes in more recent accounts such as Shapere (1982), and offering systematic arguments against it. Second, it takes the debate to a new level by challenging the epistemic distinction between direct and indirect observation in its most contextualized form, as defended by Elder (2025).

The paper proceeds as follows: Section 2 provides a conceptual basis for explaining what makes an observation direct or indirect. This will help us better understand the role of these qualifiers in scientific observations. Section 3 unfolds the historical and philosophical foundations of the distinction. Section 4 offers a series of systematically argued examples against the principled distinction based on sense perception, concluding that the traditional distinction is irrelevant to scientific observation. Section 5 focuses on more recent attempts to preserve the epistemic distinction, namely those of Shapere (1982) and Elder (2025) in a general sense and in a restricted sense, respectively. These criteria are, respectively, interference versus non-interference and intervention versus non-intervention/model dependence versus model independence. In Section 6, I critically examine the criteria used by the practice-oriented philosophers. Through a conceptual analysis that is supported by examples from practice, it is shown that these approaches face significant challenges. Some concluding remarks are given in Section 7.

2 The Conceptual Basis

In this section I shall try to lay the grounds for explaining what makes an observation direct or indirect. In the pursuit of knowledge about the natural world, human observers have historically relied on perception as the most primitive and immediate method of observing the phenomenon under study. That is, the sensory perception of the phenomenon itself (e.g., fire) was considered observation proper. Observing the traces or effects (the smoke) of the phenomenon, by contrast, was not considered observation but rather inference. Only by introducing other observational methods (such as telescopes) is it possible to use the qualifiers like "direct" and "indirect." Note that one might intuitively think, based on a philosophical position, that seeing the smoke is surely an indirect observation of the fire, but the empiricist doctrine (which is considered to be the backbone of modern science) refuses to use the concept of observation for such a case because it involves a causal inference. One might therefore be tempted to think that in order to delineate which observations are direct or indirect, one must contrast different methods of observation.

Within scientific practice, however, theoretical, social, and political dynamics may lead to the use of the qualifier "direct" and "indirect." When conflicting observations arise, scientists may use these qualifiers, not based on the observational methods, but as a rhetorical and strategic move to elevate certain observations as more authoritative. This conflict can occur between scientists who work within competing paradigms or theoretical frameworks.[1]

Therefore, the distinction between direct and indirect is either based on the observational method or the sociopolitical aspect of scientific practice. In the former case, for example, observations made with an optical telescope and with the naked eye might be called indirect and direct observations, respectively. In the latter, however, one can imagine that two contradictory observations might motivate a group of scientists to call the one that confirms their theory direct and the other indirect.

The directness and the indirectness of a method is based on the characteristics of the method of observation itself in comparison with the characteristics of another method(s). Indeed, one can mention several methodological comparisons, such as perceptual versus non-perceptual, mediated versus unmediated, visual versus statistical,

[1] For a discussion of the role of non-epistemic factors and sociopolitical arguments for direct/indirect observations of gravitational waves, see Collins (2018).

and so on.[2] Each of these observational methods, alone or as a comparative pair, has a number of characteristics based on the physical mechanism by which the observation is made, its field of application and its definition. This gives rise to different notions of directness/indirectness in different fields of observational physics. This includes, but is not limited to, the extent to which the observed and the observer are physically proximal and distal to each other. This distance-based distinction should be sufficient to understand what is meant by methodological characteristics in our case, since the arguments presented in this paper will largely focus on astronomy and astrophysics.

That considered, an observation, to wit, is not direct because it is epistemically superior, but rather it is epistemically superior because it is direct. I take it to be a mistake to say that a method of observation is direct because, for example, it is reliable. It should be said, if that were really the case, that this method is reliable because it is direct. The epistemic dimension follows from the methodological distinction. If a direct observation is considered epistemically superior to an indirect observation, it should have to do with the method of observation.

Therefore, there are two distinct ways for differentiating between direct and indirect observation:

Methodological: The distinction between direct and indirect observation arises from the two different approaches to observing a phenomenon (or two sets of phenomena), one of which is called direct and the other indirect, based on criteria such as physical proximity or the design of the observational/experimental setup.

Non-methodological: Practitioners within a given field label certain observations as direct, as opposed to others that are labeled indirect, for example, for social or political reasons.

If there is any epistemic relevance, it should arise only from the methodological one. This is because, in scientific observation, epistemic justification is tied to the reliability of the method of observation and the theories involved in its specific context. Thus, the evidence is considered reliable in a given situation when the method is tested, calibrated, cross-checked, and validated. The justification of observational evidence should not be based on social and political factors, at least in principle, as they affect the objectivity of the evidence.[3]

This implies, first, that epistemic relevance, by definition, cannot arise from the non-methodological approach. Second, there is no independent epistemic category of direct-indirect distinction. Thus, if one wants to deflate the general (principled) epistemic distinction between direct and indirect observation, as I do, one has to show that the methodological distinctions do not allow for strict epistemic distinctions. In other words, calling an observational method "direct" rather than "indirect" does not imply that the former is epistemically superior to the latter.[4] A further step is needed to explicate why the direct method is epistemically superior. This additional step should provide a clear and non-arbitrary correspondence between the methodological and epistemological differences. Without it, we will fall into epistemic relativism.

[2] Sometimes the information channel gives rise to different observational methods, such as electromagnetic channel, gravitational wave channel and so on.

[3] I am grateful to reviewer one for bringing this point to my attention for clarification.

[4] The claim does not exclude the possibility that one methodologically direct observation is epistemically superior to another methodologically indirect observation. It merely asserts that the epistemological difference does not follow from the methodological difference.

One needs to identify a non-arbitrary set of epistemic criteria or modes of justification that consistently appear in (or can be reliably related to) the characteristics of those methods which are considered direct. One way to think about this would be, for example, that direct methods of observation categorically rely on fewer sets of assumptions or inferences, and this generally results in more reliable observational outcomes. If this is not possible, as it appears to be the case, the alternative is to create a correspondence between direct methods and epistemic superiority based on arbitrary and non-objective criteria.

Having established what makes an observation direct or indirect, the subsequent section will proceed to delineate the historical and philosophical underpinnings that form the basis for a principled epistemic distinction between direct and indirect observation, with a foundation rooted in sense perception.

3 Revisiting Empiricism

There is a line of empiricism whose roots are ancient, but can be clearly traced back to British empiricism and reappeared under new guises through logical positivism to constructive empiricism (van Fraassen, 1980). The main epistemological goal, according to this line of empiricism, in maintaining sense perception in the process of observation has to do with the confidence that individual observers place on their perceptually gathered data. This implies, broadly construed, "direct" sensual contact, especially seeing, with the observed object is a necessary condition for an epistemically reliable observation.

It is important to note that the empiricist view derives from the principled distinction between the observable and the unobservable which is based on human sensory perception. According to this view, it is epistemically and ontologically privileged in a principled way to believe in the observable phenomena and not to believe in the unobservable phenomena because they are considered speculative. It should be mentioned that van Fraassen's position is relatively different, and therefore his attitude toward direct and indirect observation is correspondingly different. The literature on the distinction between observables and unobservables is enormous (van Fraassen, 1980, 2001; Muller and van Fraassen, 2008; Musgrave, 1982; Friedman, 1982; Foss, 1984; Creath, 1985; Ladyman, 2000; Fine, 2001; Teller, 2001; Muller et al., 2005). However, some clarification of the relationship of this distinction to direct and indirect observation is in order. This will help us to understand van Fraassen's position more clearly.

van Fraassen explicitly claims that,

> We can *detect* the presence of things and the occurrence of events by means of instruments. But in my book that does not generally count as observation. Observation is perception, and perception is something that is possible for us without instruments (2001, p. 154, emphasis added).

However, he has also argued that,

> A look through a telescope at the moons of Jupiter, seems to me a clear case of observation, since astronauts will no doubt be able to see them as well from close up (1980, p. 16).

Maxwell (1962) famously argued that thanks to instruments that help us in the process of observation, what was previously counted as unobservable (a theoretical entity) should now be counted as directly observable. Responding to Maxwell's argument against the importance of distinguishing between the observable and the unobservable, van Fraassen distinguishes between detection and observation, claiming that what is unobservable to humans can be detected by observation but not directly observed. van Fraassen argues that while such a distinction may lack ontological significance, it does have epistemological significance (1980, p. 19).

That considered, the theory-observation distinction for constructive empiricists seems to differ from the traditional one in that it is based on all that is in principle observable to humans, not on what has already been observed. It seems to me that by the term "observation" van Fraassen means direct observation with the naked eye and the rest is indirect observation (detection) in the sense of involving inference.

To reiterate, since the distinction is based on what is in principle observable to humans as an epistemic community, it rests on the idea that observation of something is necessarily an activity of perceiving the thing by individual scientists, and perceiving does not involve inference. Moreover, since instrumental observations rely on the theory of the instrument, the background knowledge, and the theory of what is being observed (in the case of testing predictions), which normally form the basis of the inferential bridge between the observer and the practically/in principle unobservable object (see Churchland's argument in Section 4), the exclusion of instrumental observations relies on the idea that observation is independent of theory. This makes the reasoning consistent with inference-free perception.

In short, the tenets behind this line of reasoning are:

1. Observation is individualistic.
2. Observation is perception.[5]
3. Observation is independent of theory/non-inferential.

It is important to keep in mind that the principled epistemic distinction between direct (naked-eye) and indirect (instrumental) observation relies crucially on these tenets. Moreover, the last tenet reappears under new guises in Shapere's distinction based on the notion of interference and Elder's distinction based on the notion of intervention and/or model-ladenness.

Having identified the basic tenets of the principled epistemic distinction between direct (naked eye) and indirect (instrumental) observation, in Section 4 I challenge the empiricist idea that naked-eye observation is inherently superior to instrumental observation. For that reason, I first provide reasons from scientific practice and philosophy of science. Second, I systematically argue against the principled distinction based on sense perception, concluding that the traditional distinction is irrelevant to scientific observation. In Section 5, I will resume to continue the journey of exploring the

[5]Fodor (1984) argues for a theory neutral observation by suggesting an alternative psychological account of the relation between cognition and perception. However, he admits that perceptual observation involves inferences. See also Brewer and Lambert (2001) for arguments about theory laden observation in perception and the rest of scientific practice.

basis of the distinction between direct and indirect observation by turning to practice-oriented philosophers. In particular, I will focus on examining epistemic distinctions made by Shapere (1982) in a general sense and by Elder (2025) in a restricted sense.

4 The Decentralization of Sense Perception

The first move toward decentralizing the senses involved enhancing sensory capabilities through inventions like telescopes and microscopes. Contemporary observations often combine experimentation with controlled environments, furthering the decentralization of sense perception. This approach makes rare or unnoticeable natural occurrences observable. These steps challenged conventional beliefs and necessitated reevaluating those established tenets of empiricism. Correspondingly, the concept of observation has evolved in response to scientific and technological progress, challenging empiricism. Observations of the Higgs boson and gravitational waves are evidence of this. The need for direct perceptual observation has been discarded, leading to a reconsideration of distinctions like direct and indirect. As will be shown later in this paper, in modern observations the crucial epistemological difference lies in the reliable scientific information provided, not in the directness/indirectness of the methods of acquisition. Justification depends on the reliability of the method and the theories, laws and assumptions involved, without the need for discrete qualifiers such as direct and indirect. Reliability is case-specific and there is no universal criterion for the epistemic categorization of observations. Therefore, in the realm of scientific practice, a great number of scientifically justifiable observations do not support any of these tenets either individually or in group.

In parallel, the so-called post-positivism philosophy of science (Hanson, 1958; Kuhn, 2012; Maxwell, 1962; Bogen and Woodward, 1988; Chang, 2005; Azzouni, 2004) has extensively criticized this line of empiricism especially on the level of the third tenet, namely the matters surrounding theory-laden observation.

Besides, the strict distinction between observables and unobservables made by philosophers of science is shown to be questionable. Compelling arguments against a strict distinction between observables and unobservables can be found, among others, in (Maxwell, 1962; Hacking, 1981; Musgrave, 1982; Churchland, 1985; Bogen and Woodward, 1988; Ladyman, 2000; Massimi, 2007). For example, Churchland (1985, pp. 39-41) distinguishes many modes of unobservability, including things that are very distant from us spatially or temporally and things that are too small to be observed by our senses. One might consider that the former is observable but yet unobserved to humans (practically unobservable), but the latter is unobservable in principle. He notes that for van Fraassen there is a principled distinction between things that can be observed with the unaided senses and things that are practically unobservable (like a planet in a distant galaxy) on the one hand, and things that can not be observed in principle by humans (like an electron) on the other. Churchland rightly points out that van Fraassen's principled distinction is unsatisfying because our epistemic access to both the practically unobservable and the in principle unobservable relies on ampliative inferences. So, whether it is spatial distance or spatial size, the epistemological problem is the same: bridging the gap in both cases involves "ampliative inferences and underdetermined hypotheses" (p. 40).

To further justify my critique of maintaining a principled distinction between direct and indirect observation on epistemological grounds - even for those committed to constructive empiricism - I present a series of systematically argued examples. Taken together, these examples support the view that the epistemic significance of an observation is context dependent. Moreover, they challenge the notion that naked-eye observation is inherently superior; in some cases it may even prove epistemically inferior or detrimental, depending on the specific circumstances in which the observation is made.

I begin with a story that simply offers reasons to not always rely on my senses.

One day, I opted to extend my stay at the workplace to complete reading Grover Maxwell's impressive article on "The Ontological Status of Theoretical Entities." As I navigated the final dense segments of the article, I acknowledged being unable to comprehend its depth any further. Suddenly, my gaze shifted towards the window where I spotted flickering lights in the darkness. At that moment, I decided to depart. While walking along the elongated corridor towards the nearest elevator, I observed an individual walking from the opposite direction. As I entered the elevator, I realized I was alone. The following day, I couldn't help but ponder whether I had hallucinated or if there was truly someone there. I considered the best course of action to resolve the mystery. Undoubtedly, the optimal solution is to review any available surveillance footage in the corridor. If we observe an individual in the recording, we can conclude that I had really seen a person. If not, the conclusion is that my experience was a hallucination.[6]

It could be argued that the story describes a situation in which the human senses are not in their ideal state, due to fatigue and lack of concentration, and you were not in a proper observation situation.

That being granted, this time consider a football match in which referees are observers to ensure that the rules are followed. In a particular football match, a few minutes into the game, the experienced referee, who has a clear vision, observes a foul near her position, which gives her the best view of the game. The referee decides to call the foul. However, the assistant referee in the opposite direction on the diagonal challenges the decision. To resolve the controversy, the Video Assistant Referee (VAR) should be checked. If the VAR confirms the referee's observation, it means that there was indeed a foul, otherwise the referee's observation was inaccurate.

Before drawing any conclusions from these accounts, let me situate some aspects of them in language that will enhance this article. The act of observing the individual from the opposite direction in the first case and the referee's observation in the second case are typically referred to as "direct observation," while reviewing the footage/the VAR is considered "indirect observation." The epistemic status at play is the reliability of the sensory experience compared to that of the footage/VAR.

[6]This is different from natural phenomena such as rainbows, which van Fraassen (2001) calls public hallucinations, even though they can be recorded and photographed. His rejection of the observation of rainbows stems from his denial of the possibility of observing light. His discussion of public hallucinations is irrelevant to the point I am making here.

One possible objection is that reviewing the footage/VAR is also direct as we can "directly see" the person/event. However, I argue, in order to equate this directness with the sensual experience, we must ignore the need to rely on a set of scientific theories and technologies that make this "directness" possible. Furthermore, this is an observation of the past made through an apparatus.[7] If observing a past event with a theory-laden apparatus does not preclude directness, why should an observation of a distant galaxy with a high-quality telescope be considered "indirect" when compared to naked-eye observation?[8]

One might further object that the above line of reasoning misses the mark since scientific observation is a meticulous activity, rendering the concerns in the previous story insignificant. Therefore, direct observation with the naked eye maintains its epistemic superiority. My answer would be that scientific observers predominantly favour recording the phenomenon under investigation to extract the most reliable information. This approach relies less on their senses and memory while affording them the opportunity to review the footage or recording multiple times to ensure accuracy.

To further support this claim, consider observing the Moon, once with the naked eye and once with a telescope, to study its nature and topology. In Galileo's time the latter, indirect observation, was considered unreliable, although he saw more details and gathered more information about the Moon than is possible with the naked eye. In contrast, I doubt that a single scientist in this century observes the Moon with the naked eye for scientific purposes.

This suggests two facts: first, the epistemic significance of an observation, including the reliability of the method and the inferences involved, is contextual. By contextual I mean whether the observational claim meets the epistemic standards set by the scientific community, since these standards are subject to change. Second, telescopes are now regarded as being able to collect more light from the observed object than the naked human eye, thus providing clearer images and more detailed information, i.e. a more reliable means of observation.

The lesson to be drawn from the above examples is that the distinction between direct and indirect observation seems to be epistemically irrelevant, even in cases involving naked eye observation. An additional, though controversial, lesson is that since the distinction has no epistemic bite, it might be better to drop it altogether when it comes to evaluating scientific observations. However, one might argue for retaining the distinction on methodological grounds and for heuristic reasons.

In Section 5, I will turn to epistemic distinctions between direct and indirect observation that have been proposed by practice-oriented philosophers. In particular, I will focus on examining epistemic distinctions made by Shapere (1982) in a general sense and by Elder (2025) in a restricted sense.

[7] Here I omit the fact that every observation is an observation of the past, see Kosso (1992).

[8] See Brown (1985) and Chalmers (1985) for a detailed discussion of how Galileo showed the unreliability of naked eye observations as opposed to telescopic observations.

5 The Practice-oriented Turn

5.1 Information Transmission as the Basis of Observation

Dudley Shapere's (1982) account of observation is arguably the most well-known to deviate from the empiricist view, which centers on sense perception as the foundation of observation. Shapere contrasts the narrow view of a philosopher with the permissive stance of an astrophysicist in order to challenge the dominance of empiricism in discussions of observation. He claims that the latter has a rational justification for such permissiveness (p. 511). While Shapere doesn't explicitly define the philosopher's use of "direct," he interprets it as being tied to perceptual inaccessibility. He suggests that the philosopher is an excessive empiricist who equates "direct observation" with sensory perception. Shapere rejects the philosopher's emphasis on sense perception, but retains the direct-indirect distinction on methodological and epistemological grounds. Shapere's view highlights a fundamental difference between the philosopher's perception-based genuine observation and the astrophysicist's focus on justifiable observation through transmitted information (p. 517). However, as we shall see, his view is not entirely free of empiricist motivations which regard "direct" observation as non-inferential and epistemologically superior to inferential observation.

Shapere's proposal for observation is the following:

> *x is directly observed (observable) if:*
> *(1) information is received (can be received) by an appropriate receptor; and*
> *(2) that information is (can be) transmitted directly, i.e., without interference, to the receptor from the entity x (which is the source of the information)* (1982, p. 492, emphasis in original).

The proposal shows that Shapere's understanding of observation is based on the transmission of information, so the direct-indirect distinction is based on the idea of without-with interference. Furthermore, the term "receptor" indicates that the observer need not be a human agent. That is to say, Shapere has a fundamentally different understanding of observation. His use of the term "receptor" departs from the empiricist philosopher, but he sympathizes with her by preserving the epistemic distinction. This is an incomplete revolution against the empiricist understanding of observation. This paper is an attempt to complete it.

Although Shapere departs from the empiricist on the basis of the distinction, the motivation behind it remains intact. He argues against the philosopher for conflating the problem of observation with the problem of perception (pp. 507-508), while at the same time preserving a distinction in the former that is reminiscent of the latter. The question is, what makes him fail to recognize the distinction as part of this conflation?

In his paper, Shapere considers an objection raised by a philosopher who is driven by her overly empiricist beliefs. The philosopher argues that this kind of observation, which Shapere calls direct, requires a great deal of background information and inference. Observation should be free of inference. Therefore it should not even be called observation. This is because it tends to obscure an obvious epistemic difference between non-inferential observation and inferential observation.

Shapere replies to this objection as follows: for the philosopher, he argues, inference is seen in the logical sense, so any observation that involves calculation and deduction is inferential. However, in the process of seeking knowledge, Shapere argues, one should not conceive of inferences in the logical sense as the philosopher does, but rather make a distinction between inferential and non-inferential on epistemological grounds. That is, whether we have specific reasons to doubt a conclusion or not. If the reasoning and the conclusion are subject to specific doubt, it is inferential; if we are confident about our reasoning and conclusions, it is non-inferential (p. 517).

It seems Shapere has built the distinction based on two distinct information channels (solar neutrino versus solar photons) which represent two methods, and attaches epistemic terms like confidence, doubt, specific reason for doubt, and so on to them.

As explained, Shapere abandons the two tenets of "observation is perception" and "observation is individualistic." For him, the basis of observation is the transmission of information, and this information can be captured by a receptor. However, one of the empiricist tenets seems to have implicitly persisted and reappeared under new guises in Shapere's epistemic distinction between direct and indirect observation: "direct observation is inference-free (non-inferential)." For empiricists, direct observation is non-inferential because theory and instruments are excluded from the process of observation. For Shapere, direct observation is non-inferential because there is no interference in the transmission of information between the source and the receptor, which provides confidence in the observational result.

Therefore, this time the idea rests on two foundational premises:

1. "Direct" observation is inference free (non-inferential).
2. An inference free observation is epistemically superior to an inferential observation.

With this in mind, let us now turn to the most recent attempt to make an epistemic distinction between direct and indirect observation, namely Elder (2025). Elder's distinction is explicitly contextualized to the observation of gravitational waves and possible related cases. Thus, a principled distinction is abandoned.

5.2 The Observation of Gravitational Waves

Russell A. Hulse and Joseph H. Taylor (1975) announced the discovery of the first binary pulsar system, PRS B1913+16. Subsequently, Taylor and Weisberg (1981) demonstrated the existence of gravitational waves by observing the energy decay of the binary from electromagnetic waves received by a radio telescope.

On February 11, 2016, the LIGO-Virgo collaboration announced that they have observed gravitational waves from a binary black hole merger received by the LIGO interferometer detectors. The collaboration called the observational event, GW 150914, "the first direct detection of gravitational waves and the first observation of a binary black hole merger" (Abbott et al., 2016).

Jamee Elder (2025) has a positive view of the epistemic distinction between direct and indirect observation in the case of the observation of gravitational waves.

For her, the idea behind the importance of gravitational waves interacting with the measuring device is that the data represents the properties of the gravitational waves themselves in the LIGO-Virgo case, but in the Hulse-Taylor-Weisberg case, we need to make an inference from the radio waves to the existence of the gravitational waves. However, since gravitational waves can interact with both measuring devices, that means the fundamental difference between the two cases depends on how one designs a detector: whether it is designed to detect gravitational waves or radio waves (2025, p. 9).

I think, if one is tempted to generalize this interaction-based distinction and apply it to the distinction between direct and indirect observation, one is forced to count the observation of an object with our naked eyes as indirect since our eyes do not necessarily interact with the object itself; rather, it is a chain of interaction in the sense that the photons interact with the object and eventually interact with the pigments of our eyes.

However, Elder goes beyond this naïve idea, and she considers a further reply to this interaction-based distinction in her paper, namely if one re-describes the modelling differently, it directly influences the decision one makes about whether an observation is direct or indirect. Essentially, if one counts the Hulse-Taylor binary as part of the detector, one can clearly recognize the fundamental similarities between the two cases (2025, p. 9).

Elder recognizes that there is no standard reason for dismissing this version of redescribing the situation. Moreover, the objection is pervasive in the sense that it is applicable to other examples, such as Shapere's construal and the modelling of receptor and measuring device. As a result, she draws back to build the distinction based on the notion of "intervention" between the two cases of the observation of gravitational waves.

She argues that one is justified in not including the Hulse-Taylor binary as part of the detector because it cannot be intervened upon. In contrast, the LIGO-Virgo detector on Earth can be. The intervention is important, she thinks, because it allows one to calibrate and test the detector. Therefore, she claims that the depiction of the detectors the way she argues for them is not arbitrary. She proceeds to justify the direct and indirect distinction based on the idea of intervention. The motivation behind it, I believe, comes from the idea that the detector can be seen as a "black box" (epistemically secured) as it is calibrated and tested through intervention.

This appears to be a similar strategy to Shapere's as employed in his construal of direct observation, namely attempting to provide an inference-free domain between the observed object and the receptor/detector. In Shapere's case, non-interference in the information carrier (neutrinos) leads to direct observation; in Elder's case, intervention in the detector leads to direct observation. Otherwise, her justification for excluding the Hulse-Taylor binary from the detector becomes ungrounded. The fact that both Shapere and Elder end up relying on the term "confidence" can be considered supportive of my claim (1982, p. 517; 2025, pp. 9-10).

As mentioned in the introduction, Elder emphasizes throughout the paper that direct observations are not *inherently* superior to indirect observations. Additionally, Elder stresses that what is epistemically important in the two cases of gravitational wave observation lies in how we justify our confidence in each epistemic situation, or in other words, how we delineate the nature of justification in each case. However, as will be shown in detail, there seems to be two interpretations for Elder's epistemic distinction and two criteria for the distinction. A closer look at the situation reveals that these two criteria are in tension with each other.

I conclude that there is a fundamental similarity between Shapere's and Elder's strategies of argumentation. Both begin by claiming to build on scientific practice, proceed to analyzing methodological differences (solar neutrinos versus solar photons and gravitational waves versus radio waves), and end with a contrived observational setting (based on non-interference in Shapere's case and intervention on the detector in Elder's case). Finally both retreat to purely epistemic reasons.

Next, I will take a closer look at the criteria used by these practice-oriented philosophers, Shapere and Elder. These criteria include interference versus non-interference and intervention versus non-intervention/model dependence versus model independence. Supported by examples from practice, I demonstrate that these approaches face significant challenges, both together and individually. The following Section 6 is divided into two parts: In Subsections 6.1, 6.2, and 6.3, I will primarily engage with Shapere's distinction, and in Subsection 6.4, I comment on Elder's distinction.

6 The Incongruence of the Epistemic and Methodological Distinction

6.1 Confidence Is Not a Discrete Concept

Let us assume that the distinction between direct and indirect observation seems plausible for methodological reasons, whether one's reference is perception or any other method. While physical proximity between the observed and the observer - such as directly perceiving the observed or engaging in direct physical interaction - may be advantageous, it does not necessarily guarantee the epistemic reliability of the observation. For instance, consider two scenarios: in the first, there is direct physical interaction between the observer and the observed, but understanding the interaction requires a set of theories, laws, and assumptions that are epistemically weak. In the second scenario, the observation is indirect, involving a physical distance between the observer and the observed, with additional intermediary interactions. However, in this case, the explanation relies on just one reliable physical theory or a well-supported and reliable set of theories, laws, and assumptions. In these situations, we are compelled to place more confidence in the second, indirect observation due to the stronger epistemic foundation.

For illustration, consider the following concrete example. This example is not an exact representation of the scenarios, instead it is an instance from scientific practice that provides support. As demonstrated by Hasok Chang's (2004) notion of epistemic iteration, indirect methods can become more epistemically robust than the simpler, more direct methods on which they were initially calibrated. The contrast is between human temperature perception and a thermoscope. Chang explains that confidence in thermoscopes' reliability was built by comparing their results with human temperature perception. However, it later turned out that thermoscope measurements could correct sensory perceptions (pp. 30-40). Thermoscopes were then developed into calibrated thermometers. Their calibration is based on assumptions about heat and expansion, among other things. Despite being theory-based, we now agree that thermometers are more accurate and reliable than human temperature perception.

This reasoning highlights a contradiction in Shapere's argument. On the one hand, he criticizes the empiricist philosopher for framing inference in purely logical terms, without considering whether the results inspire epistemic confidence. On the other hand, he seeks to establish an "inference-free" domain to support his notion of "directness" (as evidenced by his contrasting example in his paper). The contradiction arises from the fact that, if we are to understand "inferential" and "non-inferential" in epistemological terms, then both the physical proximity between observer and observed (methodology) and the theoretical reliability (confidence in the involved theories, laws, and assumptions) must be considered to make a robust epistemic claim. Note that I am not arguing here for the idea that observation is epistemically superior if it relies on fewer steps of theoretical reasoning, in the sense that one needs one physical law or theory to extract the information about the observed or several ones. This also seems to be a misleading idea. What I am suggesting is that regardless of the amount of theoretical reasoning involved and the nature of the methodological channel, what matters is whether or not one is confident about it in the specific context.

Moreover, since observation is a different problem from perception, and since the main purpose of observation is epistemic, as Shapere admits, one needs to provide a set of epistemic standards that correspond to, at least, a considerable number of methodologically direct or indirect observations in order to be able to maintain the distinction epistemically. This requires a set of prescribed epistemic standards to be applied to each individual observation in order to decide which group it belongs to. A vague methodological distinction might serve as a rule of thumb, but an epistemic distinction, if any, must be sharp.

It is not clear to me how one can reasonably construct a distinction between discrete qualifiers like direct and indirect on the basis of continuous and gradual epistemic terms like doubt and confidence. Unlike direct and indirect, confidence and doubt have a gradual character.[9] I argue that this very fact renders the distinction irrelevant. Furthermore, the factors that create consensus on confidence about observational results are contextual. It might not be an easy task if one attempts to find two observational results where the factors behind the consensus on the confidence in these observational results are relatively the same. Despite the fact that the criteria for doubt and confidence vary from case to case, there is an immediate response to Shapere, namely that his epistemic distinction implies that if we were equally confident about the conclusions we draw about the center of the Sun from the information we obtain from the photons of the Sun's surface (which is quite imaginable), these observations would be non-inferential and hence direct. If that is the case, the whole edifice that Shapere has built for the qualifier direct based on methodological reasons in the case of solar neutrinos collapses. This demonstrates that the epistemic distinction does not correspond to the methodological distinction. More precisely, the method of observation is neither a necessary nor a sufficient condition for epistemic confidence.

Shapere seems to reduce confidence to non-alteration. However, his focus on non-alteration of information leads to the mistaken view that it is the only reliable source for calibrating and testing observational channels. This is a red herring, as calibration and testing can be achieved through other means. Moreover, it suggests that

[9] For a Bayesian account of these gradual epistemic conditions, see Bovens and Hartmann (2004).

noninterference is the sole contributor to epistemic confidence across a wide range of observations, which is overly simplistic.

Consider Galileo's observation of the lunar mountains. At the time, the only empirical justification for the existence of these mountains came through Galileo's newly invented telescope. He observed bright and dark spots on the Moon and argued that they could be explained by the Moon being mountainous, using the Earth's sunrise phenomenon as an analogy. Galileo assumed that the Moon and the Earth shared enough similarities to make this comparison, an assumption grounded in the Copernican theory. For instance, he assumed that, like the Earth, the Moon's material reflected sunlight, which accounted for its brightness (Ariew, 1984; Spranzi, 2004). William Shea (2000) further demonstrated that Galileo's observation of the lunar mountains was not an independent test of the Copernican system but rather a consequence of his pre-existing commitment to Copernicanism during the observational process.

Furthermore, Galileo successfully calibrated and tested his telescope by observing terrestrial objects, and he distributed copies of his telescope to ensure reproducibility of his results. Nevertheless, his observations remained controversial at the time (Galilei, 2016; Franklin, 2017).

According to Shapere's criterion for direct observation - non-alteration - Galileo's observation of the lunar mountains qualifies as direct, meaning that his contemporaries should have had confidence in his observations. That is because, the photons from the lunar mountains reached Galileo's telescope without interference.

The lesson from Galileo's observations is twofold: first, his testing and calibration methods did not rely on non-interference. Second, the controversy surrounding his observations stemmed from the fact that his contemporaries did not accept the underlying assumptions and theories on which his claims were based. Thus, while Galileo's observations meet Shapere's criteria for directness, his framework does not fully capture the epistemic complexity of his work as the epistemic standards themselves were in a state of transformation.

6.2 When Does the Observational Process End?

As you can see from Shapere's proposal above, for direct observation to be non-inferential, there should be no interference in the transmission of information between the source and the receptor. The non-interference condition is only possible if one disregards the additional interactions and inferences required after the neutrinos are captured to extract the relevant information. It follows that, for Shapere, the initial interaction between the neutrinos and the receptor is the end of the observation process. This is epistemically significant for him, because non-interference keeps the information from alternating. Interestingly, in addition to the "observation is non-inferential" tenet, counting the capture of the neutrinos by the receptor as the end of observation and black-boxing the internal inferences bears similarities to the "observation is perception" tenet.[10]

Therefore, Shapere needs the following assumptions to construct the distinction:

[10] I am grateful to reviewer three for bringing to my attention the need to better clarify this point.

1. The information carrier should be some kind of physical entity: for example, neutrinos or photons.
2. The physical interaction between the carrier and the receptor is the end of the process of observation.
3. The epistemic significance of directness over indirectness comes only from the arrival of the information carrier without interference at the receptor.

When does the process of observation end? This question is crucial because it dictates one's claim about the distinction between direct and indirect observation. There does not seem to be a prescribed set of criteria for this. Again, it is case-specific, and the decision depends on the nature of the phenomenon to be observed, what we want to know about the phenomenon, the amount of information or data, both in quality and quantity, that is considered sufficient, the extent to which the background information can be used to establish the result, what justifies the results, and so on. Sometimes the scientist has to collect data for several years, and the data should go through a very complicated process of analysis. Therefore, if one wants to generalize, one simply has to assume where the process of observation ends. For this assumption to work, the receptor must also be somewhat circumscribed and defined, so that the point of interaction between the information carrier and the receptor can be determined. It would be very naive to reduce an epistemic process like observation to a mere interaction between the information carrier and the receptor. The very notion of an information carrier implies that mere interaction does not yield information, that calculations, inferences, and analysis (data processing) are necessary to decipher the information. Moreover, mere interaction without the intentional aspect of observation leads us to count every natural interaction in nature as an observation. This is related to the issue of excluding the human observer from the process of observation that Shapere promotes.

According to him, the receptor does not need to be a human observer, it can be an apparatus. The motivation behind this aligns with the idea that he defends throughout the paper, namely in modern science that observation is an epistemic process rather than perceptual and we have to dispense with the perceptual dimension of observation.

I argue that here Shapere conflates the concept of human observer with human perceiver, and he again conflates information with information carrier. A human agent might not be needed when the interaction happens between the information carrier and the receptor, a human agent might also not be needed in the process of altering the information from one form to another. That is, the human does not need to perceive the raw data and what it represents. In other words, we might be justified in excluding the human agent as a perceiver, because to perceive is not a necessary condition to acquire knowledge. However, the human agent is needed as a conceiver since to conceive is a necessary condition, in the case of non-perceptual data, to acquire knowledge. That is to say, without conceiving the received information acquiring observational knowledge is not possible. In other words, the apparatus in and by itself cannot cognitively decide on an observational result. Moreover, some observations are conducted in a collaborative group; knowledge production in such cases is an interpersonal process and contains social-epistemological dimensions. Modern observational processes are not possible without involving judgment, assessment and trust due to their complex nature. In addition to that, in the process of making an observation we have to rely on many established and reasonable background beliefs (theory ladenness). This makes the process of excluding human observer even more difficult.[11]

One possible objection is that it is at least in principle possible that a highly developed and intelligent nonhuman receptor (I name it super-intelligent receptor) will be able to carry out the whole process, namely capturing the information carrier, deciphering the information through steps like data processing, analysis, calculation, inferences and interpretations and prepare even an observational report. In this case it is hard to deny that the receptor has made an observation. This might be possible with artificial intelligence and simulations. However, I find it problematic in two aspects:

First, in this case we encounter novel epistemological problems such as to what extent the methods and tools which the super-intelligent receptor employs for data processing are epistemically transparent. Shapere has based his notion of direct observation on the idea that the alterations and inferential calculations which come after capturing the information carrier are controllable, therefore the results are reliable. However, in the case of a super-intelligent receptor, the control will be lost due to the opacity of the process. Therefore, it is not clear on what criteria the reliability will be established.

Second, it is not clear whether the super-intelligent receptor will play its role in the observational process as a tool or as a genuine observer. That is, does it assist the human observer(s) heuristically or become a partner of the human in the process of conducting science. If the former is the case, it implies that the human observer is not excluded in the process. If the latter is the case, since observation is an intentional scientific process, one is tempted to argue that the super-intelligent receptor should be a conscious being or at least possess humanly features like intellectual curiosity, pursuing intellectual goals and scientific creativity and serendipity. In that case, it seems to me we are replacing a human agent with its counterpart.

As we have seen, defining and circumscribing the receptor is a necessary condition for the direct and indirect distinction claim because various conceptions of the receptor are possible in many observational cases. That is because it affects the way one constructs the non-inferential setting.

6.3 Challenging the Intervention Criterion: A Dilemma

As we have seen, Shapere's distinction is based on non-interference with the information carrier. For him, this implies non-alteration of the information. He excludes from his account the chain of interactions taking place within the receptor, since the receptor is controllable. This allows him to constitute an inference-free domain and to make an epistemic distinction.[12]

There seem to be examples in scientific practice that are inconsistent with this heavy reliance on intervention and noninterference. I am inclined to further challenge

[11] See de Bianchi (2013) for an excellent discussion of the effect of the observer on the observed, especially in microphysical cases.

[12] It seems to me that this line of argument is at least indirectly inspired by the philosophers who made a general epistemic distinction between observation and experiment based on the criterion of manipulation, as the latter is superior to the former; see Boyd and Matthiessen (2024) and Boge (2024) for a critical discussion of such a distinction.

Shapere's distinction by providing an example, which is a proposed method for directly imaging exoplanets. The detection of exoplanets has become a prominent scientific activity since the 1990s, and thousands of exoplanets have been discovered using various detection methods such as microlensing, the transit method, and radial velocity (Mayor and Queloz, 1995; Charbonneau et al., 2002).

Recently, an interesting method called "Direct Multipixel Imaging and Spectroscopy of an Exoplanet with a Solar Gravity Lens Mission" has been proposed to detect exoplanets. In this technique, scientists aim to image exoplanets by collecting emitted light from the exoplanet itself. However, since exoplanets are not self-luminous and are very faint compared to their star companions, collecting light from them requires highly sophisticated and giant telescopes, which may be impractical considering our current scientific and technological developments. As an alternative, scientists have developed a method that makes use of the sun as a gigantic lens, exploiting the bending of light due to gravitational attraction predicted by general relativity. Without going into details, the basic idea is to position a telescope on the focus line (the resulting beam of deflected light due to gravitational lensing) around 650 to 800 astronomical units (AU) away from the sun, targeting an exoplanet. However, to image an Earth-like exoplanet, we would still need a 1.3 km telescope to resolve the entirety of the so-called "Einstein ring" in one picture. To overcome this difficulty, highly developed algorithms are needed to assemble numerous pictures and resolve the complete ring if telescopes of a couple of meters in size are employed, for example (Turyshev et al., 2020).

There seems to be a challenge for Shapere, namely whether or not to count the solar gravitational lensing (SGL) as part of the detection system. This creates a dilemma for him. On the one hand, since scientists intend to employ it as a lens, one could count it as part of the detection system. If that is the case, it goes against Shapere's idea of controlled intervention over the receptor since we cannot intervene with the Sun. In that case, one might argue that the detector cannot be seen as a "black box," and thus, it is epistemically not secure. In such a case, using the idea of control over the detection system as a ground for direct detection fails. On the other hand, if we exclude it from the detection system, then it should be considered as a separate source system, and inferences are needed from it to the target system (the exoplanet). As a result, the detection obviously becomes indirect. In conclusion, although scientists already call this proposed detection method direct, if we employ Shapere's criterion of "intervention" in this case, the detection is deemed to be indirect.

6.4 The Scope and Applicability of Elder's Distinction

In the introduction to this paper, I pointed out that Elder abandons the principled epistemic distinction between direct and indirect observation. I believe that Elder's contribution has already taken the crucial step in the right direction. In this sense, the present paper is broadly in line with Elder's position.

However, she still argues for an epistemic significance of the distinction between direct and indirect observation in the case of gravitational wave observations, and hopes that such a distinction can be applied to other cases as well. In this subsection, I take a closer look at the nature of Elder's distinction, its scope, and its basis.

As a result, I highlight a number of concerns about her moderate position. These concerns suggest that my earlier worry about the incongruity of the epistemic and methodological distinctions may extend to such restricted accounts.[13]

To begin with, it is not quite clear whether Elder's distinction is merely epistemic, merely based on scientific practice, or an attempt to account for both. In other words, it is not quite clear whether her account is normative and general, normative but specific to the observation of gravitational waves, descriptive and specific to the observation of gravitational waves, or a mixture in the sense that it adheres to epistemic criteria and at the same time resonates with scientific practice.

For example, on the one hand, Elder seems to limit her discussion of the epistemic distinction to the case of observing gravitational waves, as she argues:

> My aim is not to accurately reconstruct a single account of how terms such as "direct observation" are used across contexts (indeed, I suspect that this is an impossible task) but rather to assess the potential epistemic significance of adding terms like "direct" and "indirect"—especially the case at hand: the detection of gravitational waves (2025, p. 7).

On the other hand, Elder begins to elaborate on the possible applicability of her account to other cases, as she states:

> I expect that my analysis of direct detection in this paper extends beyond its original context. To motivate this expectation, I will briefly consider how the analysis can be applied to some related cases: two unsuccessful attempts to measure gravitational waves—by Joseph Weber, and by the BICEP experiment; and the first observation of a black hole "shadow" by the Event Horizon Telescope Collaboration (2025, p. 10).

Furthermore, I take it as self-evident that the use of the qualifiers "direct" and "indirect" by the scientists in this case study motivated Elder to write a paper on the distinction. Moreover, a quote from a scientist in the LIGO case plays a central role in Elder's distinction. The quote appeared in Collins' book, namely,

> The difference is that [...] we figured out how to build a sufficiently sensitive GW receiver and since we built it, we know exactly how it works (2025, p. 9, emphasis omitted).

It follows that her account is informed and motivated by, but perhaps not bound by, scientific practice, as her main goal is to make an epistemic distinction.

I suggest the following reading of Elder's distinction. It is descriptive in the sense that it is intended to be consistent with scientific practice, particularly in the observation of gravitational waves and other observational processes with setups similar in kind to that of gravitational wave observation. At the same time, it is normative in that its epistemic significance is grounded in a criterion, namely "intervention on the detector," which implies that it should apply to any observational setup similar in kind to gravitational wave observation.

I consider this minimal interpretation of Elder's distinction to be natural. This is because, on the one hand, I recognize the fact that the use of "direct" and "indirect" varies considerably across scientific fields. On the other hand, I have limited the expectation of the scope of applicability of the distinction to those observational setups that are similar in kind to the observation of gravitational waves, rather than

[13] The comments of reviewer three, for which I am grateful, enriched the discussion in this subsection.

to all cases. Given that she uses the phrase "related cases" in the second quote in this subsection, I believe this is what Elder has in mind as well. Otherwise, the choice of these specific examples and not others remains unexplained. In fact, the scope of "similar in kind cases" is much narrower than that of "related cases."

Below, however, I offer reasons suggesting that my earlier concern about the incongruity of epistemic and methodological distinctions extends to this minimal reading of her account. The reasons do not necessarily exclude the possibility that the LIGO-Virgo case is epistemically superior to the Hulse-Taylor-Weisberg case in one way or another. I believe, for example, that LIGO-Virgo is epistemically important in the sense that it has made it possible to reliably use gravitational waves as an observational method for future observations, thus expanding our observational horizon. However, I take the reasons to issue with any strict epistemic contrast between the two observations of gravitational waves and thus challenge Elder's construction for the basis of the epistemic distinction.

6.4.1 A Tale of Two Interpretations

As discussed earlier in this paper, Elder stresses that, in order to make an epistemic distinction, what is epistemically important in the two cases of gravitational wave observation lies in how we justify our confidence in each epistemic situation. In other words, it is a matter of how we delineate the nature of justification in each case. For example, Elder states:

> The direct/indirect distinction concerns *the nature of the justification for confidence in the measurement outcome* - in the direct case, this is based primarily on models of the measuring system, while in the indirect case it also relies on models of a separate target system (2025, p. 2, emphasis added).

However, on page 9, she argues that:

> Similarly, the radio telescopes used in the Hulse–Taylor–Weisberg detections were built for the purpose of detecting radio waves and scientists have a good understanding of how they work. *The successful use of this kind of telescope over many years allows us to be confident in the reliability of the data it provides us with. However, we cannot say the same about the Hulse–Taylor binary pulsar. Hulse and Taylor did not build this binary system and we cannot control it in any way.* Thus, the way in which we go about justifying our confidence in models of this system is different than in cases where it is possible to perform controlled interventions (2025, p. 9, emphasis added).

Considering these quotations, two interpretations of Elder's position on the nature of the epistemic distinction emerge. First, the epistemic significance of the distinction refers only to the "nature of justification" and does not imply anything about the reliability of the method of observation (i.e., it does not concern truth-conduciveness). The second reading is that the epistemic significance of the distinction refers to the "nature of justification" and involves a comparative epistemic evaluation between the two cases, which includes having more confidence in one observational method than the other (i.e., it does concern truth-conduciveness).[14]

If the first interpretation is correct, I concede this distinction, since its value is primarily methodological, in my view. My main concern is with epistemic distinctions that draw on methodological distinctions between observations to establish epistemic superiority.

If the second interpretation is correct, then it implies the following: because we can intervene in the measuring system, we are confident in the models that describe it, and therefore treat it as direct; whereas, since we cannot intervene in the separate target system, we lack confidence in its models, and therefore treat it as indirect (at least in the case study at hand). This construal is sensitive to my "confidence is not a discrete concept" argument in Subsection 6.1. Despite the fact that the criteria for doubt and confidence vary from case to case, just like Shapere's case, there is an immediate response to Elder. That is, her epistemic distinction implies that if we were equally confident about the conclusion we draw about gravitational waves from the information we obtain from the radio waves (which is quite imaginable), these observations would be direct.

In addition, as mentioned earlier, she relies on a single statement by a scientist from the LIGO collaboration who argues that "the difference is that [...] we figured out how to build a sufficiently sensitive GW receiver and since we built it, we know exactly how it works" (2025, p. 9). But apart from the fact that the scientist regards "directness" as a "red herring," the quote seems to merely indicate a methodological difference. If the second interpretation is correct, it is hard for me to see how this statement, without substantiation, leads to the conclusion that we are not confident (or even less confident) in the binary pulsar models in the case of Hulse-Taylor-Weisberg. Elder might reply that it is because of the lack of intervention in the binary pulsar, but as I will show below, it is one thing to justify the distinction on the basis of "model independence" and another to justify it on the basis of "intervention." It will also be shown that they become two competing criteria for the distinction, and that dropping one of them challenges the distinction.

6.4.2 Epistemic Interdependence

Elder's analysis seems to suggest that she assumes the detection of gravitational waves can be interpreted independently of the inference of their source system. That is, she completely separates the observation of gravitational waves from the observation of their source system not just for practical reasons, but for epistemic reasons (2025, p. 8).

However, models of the target system are necessary for the calibration and design of the detector, and for the process of data analysis. Furthermore, the design and calibration of the detector as a gravitational wave detector does not guarantee the justification of the claim that the signal is gravitational waves. The justification is based on theoretical considerations, models and descriptions of the gravitational wave

[14]The two interpretations reflect a deeper philosophical difference: the internalist and externalist approaches to epistemology. I tend to think that an adequate account of scientific justification requires elements of both. However, this debate goes beyond the scope of this paper.

sources. That is, pre-existing knowledge (empirical and theoretical) about gravitational waves and their possible sources is needed to justify the choice of the direct observation method (more on this in the next point). More specifically, in this case, the models of the source system (the templates) are indispensably and directly involved in the detection of gravitational waves. Without the extensive use of the template models, it is not possible to justify the detection of gravitational waves.

So, remember that Elder's distinction is based on the idea that in the case of Hulse-Taylor-Weisberg, one must rely on the models of the pulsar, and they are not subject to intervention, so it is indirect. In the case of LIGO-Virgo, however, the situation is no different; in order to justify the observation of gravitational waves (note that I am not referring to the observation of binary black holes), it is indispensable to rely on the models of binary black holes, which are not subject to intervention.

It is true that in order to be confident that the detection is gravitational waves, one does not need to determine the exact properties of the source system. One can recognize a kind of independence in this sense. However, there is no clear independence in terms of the possibility of justifying the observation of gravitational waves without the models of the source system.

Regardless of which interpretation of Elder's epistemic distinction one intends to endorse, this blurs the distinction. In this case, the only viable defense of the distinction appears to be methodological because it can serve as a rule of thumb.

6.4.3 An Alternative Logic of Justification

One could propose an alternative logic of justification for the two observations, LIGO-Virgo and the Hulse-Taylor-Weisberg binary pulsar, which best explains the situation given their historical and scientific contexts.

In general, theory drives those astrophysical observations which are aimed to test the predictions of the theory in question. Since the early 1970s, there has been a controversy about the confidence in the correctness of the quadrupole formula (especially for astrophysical systems). Therefore, the theoretical basis for observing gravitational waves from a strong field was shaky. Before the discovery of the binary pulsar PSR 1913+16 by Russell Hulse and Joseph Taylor in 1975, all tests of general relativity were based on the first-order corrections to Newtonian theory. So, it was still not clear whether the newly discovered pulsar system would be a reliable method to measure orbital decay caused by gravitational wave emission.

By the early 1980s, however, the consistent work of Hulse, Taylor, and others had accumulated a fair amount of empirical and theoretical knowledge about pulsars, and many pulsars had been discovered. So, it was perfectly justifiable to make reliable additional inferences from them. That is, to incorporate binary pulsars into observational methods and use them as observational tools. This is exactly what Taylor and Weisberg did in their 1981 demonstration of the existence of gravitational waves.

To use a direct method in the sense of a method like the LIGO-Virgo detectors, one would have to rely on a considerable amount of theoretical description of gravitational waves and their sources, which was not available. In other words, given the reliable and established knowledge they had about binary pulsars and given the lack of detailed theoretical understanding of the properties and behavior of gravitational waves, it was more justifiable to use the indirect method in the sense of using binary pulsars to infer gravitational waves. The thin concept of gravitational waves (fewer constraints on their properties) made it more justifiable to use the indirect method.

Therefore, it was more plausible to use radio telescopes as an indirect method to infer the existence of gravitational waves. The situation was different in the case of LIGO-Virgo. The detailed theoretical understanding of the sources of gravitational waves, numerical simulations and other modeling strategies made it reasonable and justifiable to build a gigantic detector to detect gravitational waves.

This means that the indirect method (the binary pulsar plus radio telescope) was used because its reliability was already established, and the direct method was considered unreliable due to the lack of pre-existing theoretical knowledge about the observed (GW). However, it was the remarkable previous success of general relativity, detailed theoretical predictions by modeling the sources of gravitational waves, which helped establish the reliability of the LIGO-Virgo method.[15]

This alternative logic of justification between the two observations refrains from making *any* comparative epistemic distinctions between them from a retrospective point of view, considering both interpretations I provided. Rather, it examines each epistemic situation within its historical, scientific, and technological confines.[16]

6.4.4 Two Competing Criteria

As mentioned earlier, Elder sometimes draws on the notion of "model independence" to make the distinction, and sometimes she draws on the notion of "intervention."

Consider, for example, the following two quotes:

> While there may be an element of convention in the choice to treat the telescope, but not the Hulse–Taylor binary, as a detector, this choice is not arbitrary. Rather, it is based on the *interventions* (e.g., calibration and coherence testing) that we can perform to justify our confidence in treating its output as representative of some parameter of interest in the target system (2025, p. 9, emphasis added).

> Thus, the epistemic benefits of being a direct detection as stated above do not apply; when we focus on the role of the LIGO-Virgo interferometers in observing compact binary mergers all of the relevant inferences must be justified using *models* of a distant target system (2025, p. 10, emphasis added).

Here she considers the observation of a binary black hole (GW150914) to be indirect because of model dependence, but she considers the Hulse-Taylor-Weisberg case to be indirect because of the lack of intervention on the binary pulsar.

However, it is one thing to say that the observation is indirect because there is an intermediate binary system that is not subject to intervention, and another thing to say that the observation is indirect because the models of the target system being observed are assumed/involved.

If the epistemic distinction is merely concerned with the "nature of justification" and this is not arbitrary because it relies on "intervention," then the observation of

[15] See Ahmed (2025) for details about the methodological framework for inferring the source of gravitational waves and the indispensable reliance on meta-empirical observations of the general relativistic framework to overcome the challenge of underdetermination of the source.

[16] For the historical details of this subsection I have relied on Kennefick (2007).

a binary black hole (GW150914) should be considered “direct” because the black hole merger is the observed target system, not an intermediate one. However, Elder considers it “indirect” based on the idea of “model dependence.”

Gravitational waves do indeed play the role of messenger between the detector and the binary black hole. However, note that if one denies the direct observation of binary black holes simply because gravitational waves act as messengers, then one should also deny the direct observation of the binary pulsar (not GW) by Hulse and Taylor because radio waves act as messengers. Even worse, one would be deemed to deny naked-eye observations as direct because photons mediate. I assume that we want to avoid such a conclusion. Thus, the observation of binary black holes becomes indirect only when the criterion of “intervention” is dropped and “model-dependence” adopted. However, if intervention is dropped and model-dependence is instead used to define “indirectness,” then nearly all observations become indirect, and the distinction loses epistemic clarity.

Moreover, remember that Elder considers cases where “data is about visibility” to be direct observations:

> Even in complex cases involving multiple detectors and measurements, we call a detection “direct” when the data are data about the entity we are trying to detect — in the LIGO-Virgo case, strain data. In these cases, confidence in the detector combined with confidence in the data analysis methods, explains the confidence that scientists have in the detection (2025, p. 10).

Take another look at the exoplanet example I presented in 6.3. As discussed earlier, the data in the exoplanet example is about visibility, and it is a representation of the source. It is true that resolving the entire “Einstein ring” in one image requires sophisticated algorithms and relies on the Sun as a giant lens. However, there is no need to assume the theoretical model of the source (the exoplanet). It is also true that we need to make inferences about the intermediate system (solar lensing), but these inferences are not based on theoretical models, and furthermore, models of solar lensing are not models of the source, they are based on robust empirical facts about gravitational lensing. It follows that using Elder’s “model independence” criterion, the detection is direct. However, as discussed earlier, the detection is indirect using Elder’s “intervention” criterion. This is because we cannot intervene in solar gravitational lensing (SGL) to include it in the detection system.

According to Elder’s account, it does not seem possible to drop one of these two competing criteria, namely model independence and intervention. Each is needed to account for different aspects of the distinction. The intervention criterion is needed for the claim that the definition of the detector is not arbitrary (methodology), while model independence is needed for the epistemic claim (epistemology). This analysis suggests a tension between the epistemic and methodological contrasts in a single case study.

7 Conclusion

The scientific and philosophical literature on observation is replete with the use of the qualifiers direct and indirect to distinguish observations methodologically and epistemologically. However, scientific and technological developments have increasingly forced proponents of the distinction to reconsider and reformulate the basic tenets of the distinction under new guises.

The main idea I have articulated in this paper is that perhaps this time we need to reconsider the distinction altogether, without beginning to think about how to reformulate the tenets. This seems to be helpful in understanding that even if the distinction is methodologically defensible, it seems to have no epistemic bite, and even to be alien to the actual epistemic justification of scientific observations. It might be better to focus on the epistemic dimension of the concept of observation. This requires an analysis of the context and circumstances in which the observation is made in order to determine its epistemic significance.

Acknowledgements. I would like to thank the members of the "Epistemology of Modern Physics" online group, Sam Fletcher, Gregor Schiemann and Jeremias Düring for their valuable comments and discussions on earlier drafts of this paper. I thank Kimberly Taynton for her assistance in improving the manuscript's language. Versions of this paper have been presented at "The British Society for the Philosophy of science 2023 Annual Conference," the "4th Lisbon International Conference on Philosophy of Science," and the "9th biennial meeting of the European Philosophy of Science Association," from which the paper has benefited. Special thanks to Radin Dardashti for discussions and detailed comments on earlier drafts, and two anonymous reviewers for valuable comments.

Declarations

- **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., R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, et al. 2016. Observation of gravitational waves from a binary black hole merger. *Physical review letters 116*(6): 061102 .

Ahmed, S. 2025. Inference to the source. *Synthese 205*(192): 1–28. https://doi.org/10.1007/s11229-025-05015-5 .

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 .

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 .

Boge, F.J. 2024. Re-assessing the experiment / observation-divide. *Philosophy of Science 0*(0): 1–18. https://doi.org/10.1017/psa.2024.23 .

Bogen, J. and J. Woodward. 1988. Saving the phenomena. *The philosophical review 97*(3): 303–352 .

Bovens, L. and S. Hartmann. 2004, 01. *Bayesian Epistemology*. Oxford University Press.

Boyd, N.M. and D. Matthiessen. 2024. Observations, experiments, and arguments for epistemic superiority in scientific methodology. *Philosophy of Science 91*(1): 111–131. https://doi.org/10.1017/psa.2023.101 .

Brewer, W.F. and B.L. Lambert. 2001. The theory-ladenness of observation and the theory-ladenness of the rest of the scientific process. *Philosophy of Science 68*(S3): S176–S186 .

Brown, H.I. 1985. Galileo on the telescope and the eye. *Journal of the History of Ideas 46*(4): 487. https://doi.org/10.2307/2709541 .

Chalmers, A. 1985. Galileo's telescopic observations of venus and mars. *British Journal for the Philosophy of Science 36*(2): 175–184. https://doi.org/10.1093/bjps/36.2.175 .

Chang, H. 2004. *Inventing Temperature: Measurement and Scientific Progress*. New York, US: OUP Usa.

Chang, H. 2005. A case for old-fashioned observability, and a reconstructed constructive empiricism. *Philosophy of Science 72*(5): 876–887 .

Charbonneau, D., T.M. Brown, R.W. Noyes, and R.L. Gilliland. 2002, March. Detection of an Extrasolar Planet Atmosphere. *The Astrophysical Journal, 568*(1): 377–384. https://doi.org/10.1086/338770. arXiv:astro-ph/0111544 [astro-ph].

Churchland, P.M. 1985. The ontological status of observables: In praise of the superempirical virtues, In *Images of Science: Essays on Realism and Empiricism*, eds. Churchland, P.M. and C.A. Hooker, 35–47. University of Chicago Press.

Collins, H. 2018. *Gravity's kiss: The detection of gravitational waves*. MIT Press.

Collmar, W., N. Straumann, S.K. Chakrabarti, G. 't Hooft, E. Seidel, and W. Israel 1998. Panel discussion: The definitive proofs of the existence of black holes. In F. W. Hehl, C. Kiefer, and R. J. Metzler (Eds.), *Black Holes: Theory and Observation*, Berlin, Heidelberg, pp. 481–489. Springer Berlin Heidelberg.

Creath, R. 1985. Taking theories seriously. *Synthese* 62: 317–345 .

de Bianchi, M.S. 2013. The observer effect. *Foundations of Science 18*(2): 213–243. https://doi.org/10.1007/s10699-012-9298-3 .

Doboszewski, J. and D. Lehmkuhl. 2023. On the epistemology of observational black hole astrophysics, In *Philosophy of Astrophysics: Stars, Simulations, and the Struggle to Determine What is Out There*, eds. Boyd, N.M., S.D. Baerdemaeker, K. Heng, and V. Matarese, 225–2147483647. Springer Verlag.

Eckart, A., A. Hüttemann, C. Kiefer, S. Britzen, M. Zajaček, C. Lämmerzahl, M. Stöckler, M. Valencia-S., V. Karas, and M. García-Marín. 2017. The milky way?s supermassive black hole: How good a case is it?: A challenge for astrophysics & philosophy of science. *Foundations of Physics 47*(5): 553–624. https://doi.org/10.1007/s10701-017-0079-2 .

Elder, J. 2025. On the “direct detection” of gravitational waves. *Studies in History and Philosophy of Science Part A 110*(C): 1–12. https://doi.org/10.1016/j.shpsa.2025.01.002 .

Fine, A. 2001. The scientific image twenty years later. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition 106*(1/2): 107–122 .

Fodor, J. 1984. Observation reconsidered. *Philosophy of science 51*(1): 23–43 .

Foss, J. 1984. On accepting van fraassen’s image of science. *Philosophy of Science 51*(1): 79–92 .

Franklin, A.D. 2017. Is Seeing Believing?: Observation in Physics. *Phys. Perspect. 19*(4): 321–423. https://doi.org/10.1007/s00016-017-0210-y .

Friedman, M. 1982. The scientific image by bas c. van fraassen. *Journal of Philosophy 79*(5): 274–283. https://doi.org/10.5840/jphil198279536 .

Galilei, G. 2016. *Sidereus Nuncius, or The Sidereal Messenger.* Chicago: University of Chicago Press.

Hacking, I. 1981. Do we see through a microscope? *Pacific Philosophical Quarterly 62*(4): 305–322. https://doi.org/10.1111/j.1468-0114.1981.tb00070.x .

Hacking, I. 1989. Extragalactic reality: The case of gravitational lensing. *Philosophy of Science 56*(4): 555–581. https://doi.org/10.1086/289514 .

Hanson, N.R. 1958. *Patterns of Discovery.* Cambridge [Eng.]: University Press.

Hulse, R.A. and J.H. Taylor. 1975. Discovery of a pulsar in a binary system. *The Astrophysical Journal* 195: L51–L53 .

Kennefick, D. 2007. *Traveling at the Speed of Thought: Einstein and the Quest for Gravitational Waves*. Princeton University Press.

Kosso, P. 1992. Observing the past. *History and Theory 31*(1): 21–36. https://doi.org/10.2307/2505606 .

Kuhn, T.S. 2012. *The structure of scientific revolutions*. University of Chicago press.

Ladyman, J. 2000. What's really wrong with constructive empiricism? van fraassen and the metaphysics of modality. *The British Journal for the Philosophy of Science 51*(4): 837–856 .

Massimi, M. 2007. Saving unobservable phenomena. *British Journal for the Philosophy of Science 58*(2): 235–262. https://doi.org/10.1093/bjps/axm013 .

Maxwell, G. 1962. The ontological status of theoretical entities, In *Scientific Explanation, Space, and Time: Minnesota Studies in the Philosophy of Science*, eds. Feigl, H. and G. Maxwell, 181–192. University of Minnesota Press.

Mayor, M. and D. Queloz. 1995, November. A Jupiter-mass companion to a solar-type star. *Nature 378*(6555): 355–359. https://doi.org/10.1038/378355a0 .

Muller, F.A. et al. 2005. The deep black sea: observability and modality afloat. *The British journal for the philosophy of science* 56: 1–99 .

Muller, F.A. and B.C. van Fraassen. 2008. How to talk about unobservables. *Analysis 68*(3): 197–205 .

Musgrave, A. 1982. Constructive empiricism versus scientific realism. *Philosophical Quarterly 32*(128): 262. https://doi.org/10.2307/2219327 .

Shapere, D. 1982. The concept of observation in science and philosophy. *Philosophy of science 49*(4): 485–525 .

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.

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 .

Teller, P. 2001. Whither constructive empiricism? *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition 106*(1/2): 123–150 .

Turyshev, S.G. et al. 2020. Direct multipixel imaging and spectroscopy of an exoplanet with a solar gravity lens mission. https://arxiv.org/abs/2002.11871.

van Fraassen, B.C. 1980. *The scientific image.* Oxford University Press.

van Fraassen, B.C. 2001. Constructive empiricism now. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition 106*(1/2): 151–170 .

Weisberg, J.M. and J.H. Taylor. 1981. Gravitational radiation from an orbiting pulsar. *General Relativity and Gravitation* 13: 1–6 .