Philosophy of Science (2023), **90**, 629–647
doi:10.1017/psa.2023.50

PSA PHILOSOPHY OF SCIENCE ASSOCIATION
PHILOSOPHY OF SCIENCE

ARTICLE

Robustness and Dark-Matter Observation

Antonis Antoniou[1,2]

[1]Department of Philosophy, University of Bristol, Bristol, UK and [2]Department of History and Philosophy of Science, National and Kapodistrian University of Athens, Greece
Emails: aa17779@bristol.ac.uk, antantoniou@phs.uoa.gr

(Received 29 September 2021; revised 08 August 2022; accepted 09 February 2023; first published online 27 March 2023)

Abstract

Current cosmological observations place little constraints on the nature of dark matter, allowing the development of a large number of models and various methods for probing their properties, which seem to provide ideal grounds for the employment of robustness arguments. In this article, the extent to which such arguments can be used to overcome various methodological and theoretical challenges is examined. The conclusion is that although robustness arguments have a limited scope in the context of dark-matter research, they can still be used to increase the scientists' confidence about the properties of specific models.

1. Introduction

According to the current received view in cosmology, the ΛCDM model, more than 95% of the observable universe is "dark," consisting primarily of dark energy ($\approx$68%) and dark matter ($\approx$27%). However, despite the fact that the systematic research of dark matter has a history of almost 50 years, its exact nature still remains elusive, mainly due to the severe underdetermination of viable dark-matter models by the available evidence. If it exists, dark matter could be *anything*, insofar as it satisfies a minimum set of constraints based on current cosmological observations.

Martens (2022) aptly describes this minimum set of constraints as the *thin common core concept of dark matter*: assuming general relativity, dark matter is a massive field that contributes $\approx$ 27% to the current total cosmic mass-energy budget; it primarily interacts with baryonic matter via the gravitational force; and if it is a particle, its mass is expected to be between 10^{-3} and 10^{7} eV. Hence, assuming that one wishes to maintain general relativity at low accelerations, the thin common core concept of dark matter embodies the minimum set of constraints that every candidate model of dark matter needs to satisfy in order to be compatible with current cosmological evidence related to various "dark phenomena."[1] The thin common core concept of

---

[1] The term *dark phenomena* is borrowed from Martens and Lehmkuhl (2020) and refers to the various astrophysical phenomena that either contradict the gravitational laws of general relativity or require the

dark matter is therefore "common" because it is shared by all possible models for the nature of dark matter, and it is "thin" because it places very few constraints on the exact nature of dark matter. At the same time, it also reflects the progress in the field of dark-matter research in that the enrichment of this common core concept via the derivation of model-independent properties of dark matter automatically leads to a stricter set of constraints, thus leading to the reduction of viable candidate models.

As expected, the fact that the nature of dark matter is underdetermined by the available evidence has naturally led to the development of a large number of diverse, but nonetheless viable, models for dark matter. This proliferation of models has in turn led to the development of a variety of methods for the possible observation of dark matter and the probing of its properties. At first glance, the variety of models and methods of observation in dark-matter research seems to provide a fruitful ground for robustness arguments, where *robustness* is defined as "the state in which a hypothesis is supported by evidence from multiple techniques with independent background assumptions" (Stegenga, 2009). Hence, given that different models for dark matter embody different hypotheses about the nature of dark matter, the fact that there are different methods for probing the properties of dark matter can be used for the validation or the invalidation of these hypotheses.

The central aim of the present article is to examine the extent to which robustness arguments from the variability of experiments can be used within the scientific context of dark-matter research in order to overcome various methodological and theoretical challenges. The main conclusion is that although these arguments have a limited scope, they can still be used in dark-matter research to increase the scientists' confidence about the constraints on the parameter space of certain model scenarios. In particular, it will be argued that robustness arguments cannot be used for establishing which model (or models) actually captures the true particle nature of dark matter, but they can—and should—be used to increase the reliability of the results concerning the submodels of model scenarios that can be probed by more than one method.

The present article falls within the recently rising literature on the philosophy of dark matter (e.g., Vanderburgh 2014; Kosso 2013; Massimi 2018; Weisberg et al. 2018; de Swart 2020; Martens and Lehmkuhl 2020; Smeenk 2020; De Baerdemaeker 2021; Jacquart 2021b; De Baerdemaeker and Boyd 2020), and its broader aim is to explore an existing methodological challenge in cosmology regarding the proliferation of viable models for the nature of dark matter by embedding it in the current philosophical literature on robustness. Although the main conclusion partly concerns the limits of robustness arguments in dark-matter research, the present study, at the same time highlights the need for a common ground of reference in dark-matter research for the integration of results from different methods and should not be seen as introducing any kind of pessimism or skepticism about dark-matter research. If anything, it showcases how a collective and rigorous conceptual and methodological analysis of the scientific practice in dark-matter research by philosophers of science and physicists could eventually shed light on various methodological conundrums in this field.

---

postulation of additional "invisible" dark matter that causes the formation of some large-scale cosmological structures due to its gravitational pull. Examples of such phenomena are the mass discrepancies in the Coma cluster and the flat rotation curves of nearby galaxies. For a recent review of the observational evidence for dark matter and dark energy based on dark phenomena, see Jacquart (2021a).

These observations will become clearer by elucidating the existing methods of dark-matter observation and providing a possible strategy to evaluate their epistemic strength. In what follows, a brief discussion of the literature on robustness will be presented in order to identify the type of robustness analysis that is most relevant for the purposes of this article (section 2). The discussion will proceed with a presentation of the viable models of particle dark matter and the possible methods for probing their properties according to the existing scientific literature (section 3). In section 4, an analysis of the three key epistemic virtues of informativeness, model sensitivity, and reliability will be given in order to support the main argument about the limits of robustness analysis from the variability of experiments that follows in section 5.

2. Robustness and the variability of experiments

Robustness arguments have been widely discussed by philosophers of science in various contexts.[2] For the purposes of this article, *robustness from the variability of experiments* will be understood mainly in terms of what Woodward (2006, 233–35) calls "measurement robustness" and what Franklin (1989) discusses as the variability of experiments. The rationale behind Woodward's measurement robustness is that if different measurement procedures of a physical quantity that are in some sense independent of each other produce (nearly) the same result, then the result is said to be *robust* and can be used as grounds for increasing our confidence that the quantity has been measured accurately. This is because it is very unlikely that each procedure is subject to exactly the same kinds of error that would give rise to the same result, so we have good reason to believe that the result of the different measurements is reliable.

Woodward's characterization of measurement robustness is closely related to what Franklin (1989, ch. 6) describes as the *variability of experiments* in his discussion of possible strategies to justify the reliability of experimental results in high-energy physics. For Franklin, a potential agreement between the experimental results of different experimental methods automatically increases our confidence in the reliability of the results because "it would be a preposterous coincidence if the same patterns were produced by two totally different kinds of physical systems."[3] In a similar spirit to Woodward and Franklin, Stegenga (2009) provides a wider definition of robustness as "the state in which a hypothesis is supported by evidence from multiple techniques with independent background assumptions" (651). Again, the main idea behind Stegenga's definition is that hypotheses are better corroborated by results coming from multiple and independent experimental methods compared to hypotheses that are supported by the results of a single experimental method.

Given these definitions of robustness, and keeping in mind the experimental practice in dark-matter research, which is largely about constraining the parameter space of dark-matter models, what we shall call *robustness from the variability of experiments* will be understood as follows:

> Compared to hypotheses supported by only one experimental method, a hypothesis H_{DM} concerning the nature of dark matter is more robust—and thus

---

[2] See, for instance, Franklin (1989), Weisberg (2006), Woodward (2006), Parker (2011), Lloyd (2015), Basso (2017), Lisciandra (2017), and Schupbach (2018).

[3] Hacking (1983), as quoted by Franklin (1989, 166).

enjoys a higher degree of confidence from the scientific community—if it is supported by the results of two or more experimental methods.

A hypothesis H_{DM} concerning the nature of dark matter can be either about the value of a quantity of a physical property of dark matter (e.g., the cosmic relic density of dark matter) or—as is more often the case in dark-matter research—about a constraint (or set of constraints) on the parameter space of a dark-matter model regarding, for instance, the mass of a weakly interacting dark-matter particle, the couplings of mediator particles to dark matter, and so on. On the face of it, the situation in dark-matter research in which one finds a number of different methods for constraining the parameter spaces of those models provides an ideal setup for the employment of robustness arguments. The question that arises—and that is the main focus of this article—is the following: How exactly can we take advantage of the fact that there are different methods available for probing the properties of dark matter in order to increase our understanding and confidence about the properties of dark matter? Prima facie, it seems that there are two possible ways of employing robustness arguments in dark-matter research, with each one corresponding to the following two aims:

1. To reduce the number of viable models in order to decide which of the existing models actually captures the true nature of dark matter
2. To increase our confidence about the properties of a specific model of dark matter

The first aim is ultimately the most important. At the end of the day, what we are interested in is figuring out which of the available candidate models of dark matter actually captures the nature of dark matter. Hence, given that each model incorporates different hypotheses of dark matter, those hypotheses that will ultimately be supported by the results of more than one method will become more robust compared to those hypotheses that are supported only by one method—or, even worse, by no method at all.

The second aim is more modest. Instead of comparing hypotheses from different models of dark matter, the objective is to employ the results of as many methods as possible in order to increase our confidence about the allowed parameter space of a particular model. To put it simply, this would amount, for instance, to constraining the mass of a weakly interactive massive particle (WIMP) not only by using the results from direct searches but also by taking into consideration the results from indirect searches.

In what follows, it will be argued that robustness arguments from the variability of experiments, unfortunately, do not succeed in offering a possible strategy for the accomplishment of the first—and most important—aim, but they do, nonetheless, succeed in fulfilling the second aim. That is, robustness arguments from the variability of experiments are unable to provide stronger grounds for deciding between competing models of dark matter, but at the same time, they can—and should—be employed for the fulfillment of the second aim stated, namely, for increasing the reliability of results concerning the allowed parameter space of specific models via complementarity and compatibility studies.

In order to support this claim, we shall first discuss the current situation in dark-matter research by presenting some of the competing candidate models of dark matter and the possible methods for probing their parameter space.

3. Many models, four methods

As already mentioned, the fact that the nature of dark matter is underdetermined by the available cosmological evidence has naturally led to the development of a large number of diverse models regarding the nature of dark matter. Some of these models concern the particle nature of dark matter, whereas others describe the large-scale structure of dark matter in terms of massive compact halo objects (MACHOs), which may or may not be composed of baryonic matter (e.g., primordial black holes and neutron stars). For the purposes of this article, we shall restrict ourselves to models regarding the particle nature of dark matter for reasons of consistency and uniformity, as these models are already abundant enough to be used for robustness arguments and are easier to compare with each other.

Arguably, the most popular candidate for the particle nature of dark matter comes from the models of dark matter as WIMPs (Roszkowski et al., 2018), followed, perhaps, by models of dark matter as axions (Duffy and Van Bibber, 2009) and sterile neutrinos (Boyarsky et al., 2019). Roughly speaking, *WIMP* is an umbrella term referring to a large class of models, including particles that interact with ordinary baryonic dark matter via gravity and a nonvanishing force that is either weaker or at least as weak as the weak nuclear force. Axions are hypothetical elementary particles that were initially introduced as a possible solution to the strong CP problem in particle physics, and finally, sterile neutrinos are similar to standard model neutrinos, except that they have right-handed chirality and do not interact with baryonic matter via any of the fundamental interactions apart from gravity.

However, these three models do not exhaust, by any means, the available options for the nature of dark-matter particles. As mentioned in the introduction, assuming standard gravitational laws, the minimum set of constraints as stated in the thin common core concept of dark matter also allows for the possibility that dark-matter particles are made of collisionless particles, strictly self-interacting particles, or even an entire hidden dark sector of numerous fundamental dark-matter particles.[4] Given that each one of these models comes in various versions (especially in the case of WIMPs), for the remainder of this article, we shall refer to these broad classes of models as the possible *model scenarios* of dark matter, whereas the various specific examples of a particular model scenario—for example, the lightest supersymmetric

---

[4] Just to provide a sense of how long the list of candidate models of dark-matter particles is, some further proposed candidates are gravitinos, axinos, light-scalar dark matter, dark matter from little Higgs models, wimpzillas, Q-balls, mirror particles, charged massive particles (CHAMPs), strongly interacting massive particles (SIMPs), D-matter, cryptons, superweakly interacting dark matter, brane world dark matter, and heavy fourth-generation neutrinos. For a comprehensive review of dark-matter research containing useful information on various dark-matter candidates, see Bertone et al. (2005). For a review of candidates for dark-matter particles, see Feng (2010).

particle (LSP) and lightest Kaluza-Klein particle (LKP) examples in WIMPs—will be referred to as the *submodels* of a model scenario.[5]

As one might expect, the abundance and diversity of viable model scenarios for dark matter have naturally led to the development of a variety of methods for the possible observation of dark matter because each model scenario is built on a number of different assumptions and requires different experimental setups to be tested.[6] According to current scientific practice, dark matter can be observed via four different methods based on the physical phenomena on which they rely: (a) via cosmological observations, (b) directly, (c) indirectly, and (d) in collider searches. As a precaution, it should be noted that regarding methods (b), (c), and (d), the choice of labeling them as direct, indirect, and collider based reflects the relevant terminology in current scientific literature, which clearly distinguishes between these three different methods based on their methodology. However, although the methodology of each method is clearly different, whether these methods are indeed direct or indirect, and the interesting question of whether collider searches provide a possible direct or indirect way of observing dark matter, is a separate and very interesting issue that lies outside the scope of this article but nonetheless deserves to be studied further on its own merit. Finally, let us also note that although the observation of dark matter based on cosmological observations is often considered as providing strong evidence for its existence, the remaining three methods have not provided any positive results so far, other than various sets of constraints on the parameter space of some models. A brief description of the four possible methods of dark-matter observation follows.

Cosmological Observations. Cosmological observations of dark matter can be divided into two categories: (i) precision measurements of cosmological observables and (ii) observations based on the gravitational effects of dark matter on large-scale structures. Precision measurements are typically related to the nongravitational effects of dark matter on large-scale structures and the thermal history of the universe as captured by various cosmological observables related to the cosmic microwave background (CMB)—such as spectral distortions and temperature anisotropies—and to data from distance measurements of type Ia supernovae (SN Ia) and baryon acoustic oscillations (BAOs). Cosmological observations of dark matter via its gravitational effects are typically related—as the name suggests—to the purely gravitational effects of dark matter captured, for instance, by measurements of mass discrepancies in galaxy clusters, galaxy rotation curves, and gravitational lensing.[7]

Direct Detection Searches. Direct detection methods are earth-based experiments based on the interaction of dark-matter particles with ordinary baryonic matter. The basic idea behind direct searches is that if the galaxy is full of dark-matter particles (e.g., WIMPS, axions, etc.) that interact weakly with baryonic matter, then a significant amount of them will travel through Earth, enabling us to search for the interaction of

---

[5] The LSP (also known as the *neutralino*) and the LKP are probably the two most well-studied submodels for WIMPs, coming from supersymmetry and theories of extra dimensions, respectively.

[6] Following Shapere (1982), *observation* is understood here in the broad scientific meaning as an act of obtaining information about the properties of one or more physical entities via any kind of interaction that involves the communication of information from the target of the observation to the observer.

[7] For more detailed reviews on cosmological observations of dark matter, see Luković et al. (2014) and Gluscevic et al. (2019). For a review of dark-matter gravitational lensing, see Massey et al. (2010).

these particles with standard model particles by recording the recoil energy of nuclei as dark-matter particles scatter off them. Because the interaction of dark-matter particles with the nuclei is expected, by definition, to be extremely weak, direct search experiments take place in ultra-sensitive low-background experiments that are often placed well below Earth's surface in order to block out spurious particles.[8]

Indirect Detection Searches. Indirect searches for dark matter are based on the astronomical observation of the standard model particles that are most likely to be the products of the decay or annihilation of dark matter in the universe. These searches are based on the assumption that the final states of dark-matter annihilation/decay are either standard model particles of any kind (insofar as they are kinematically accessible) or unknown particles that then decay to standard model particles. Current experiments for the indirect detection of dark matter are mainly focused on the detection of three different products: (i) gamma rays, (ii) neutrinos, and (iii) cosmic rays. These three types of radiation are used for the indirect detection of dark matter for a number of theoretical and practical reasons, such as the fact that the mass scale of WIMPs in the most promising models implies that a large fraction of the generated emission from dark-matter annihilation/decay ends up in gamma-ray energies.[9]

Collider Searches. The main idea behind collider searches for dark matter is that high-energy collisions of standard model particles, such as the ones taking place at the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN), can lead to the direct pair production of dark-matter particles. If the mass of dark-matter particles is comparable to the electroweak scale, the LHC might also produce large quantities of dark-matter particles via the decay of the heavier states that are instantaneously created in high-energy proton-proton collisions. Collider searches are thus based on the possible interactions of dark-matter particles with standard model quarks and gluons that lead to a missing energy (or missing momentum) signature in the final states due to the lack of interaction between dark matter and the material of the detectors and/or the detection of unexpected particle products. There are five main processes in which dark-matter pair production could occur at the LHC: (i) mono-jet, (ii) mono-V, (iii) mono-Higgs, (iv) dark matter with top quarks, and (v) invisible Higgs decay. In mono-jet processes, the dark-matter particles are produced in association with one or more quantum chronodynamic (QCD) jets; in mono-V, they are produced in association with a vector boson; and so on.[10]

The four different methods of dark-matter observation just described and the relevant underlying physical phenomena are summarized in figure 1, with precision measurements of cosmological observables and dark-matter observation via its gravitational effects illustrated separately for reasons of clarity.[11] Each one of these methods can be used to probe and constrain various properties of dark matter, and hence, they are, in principle, valuable tools in the hands of scientists for performing

---

[8] See Schumann (2019) for a review of direct detection searches on dark matter.

[9] See Gaskins (2016) for a detailed review of indirect searches for dark matter.

[10] For a detailed review of LHC dark-matter searches, see Kahlhoefer (2017).

[11] It should be noted that this rather helpful depiction of the different methods of dark-matter observation is not exhaustive of all dark-matter research, especially with respect to the level of the phenomena that are responsible for each type of observation.

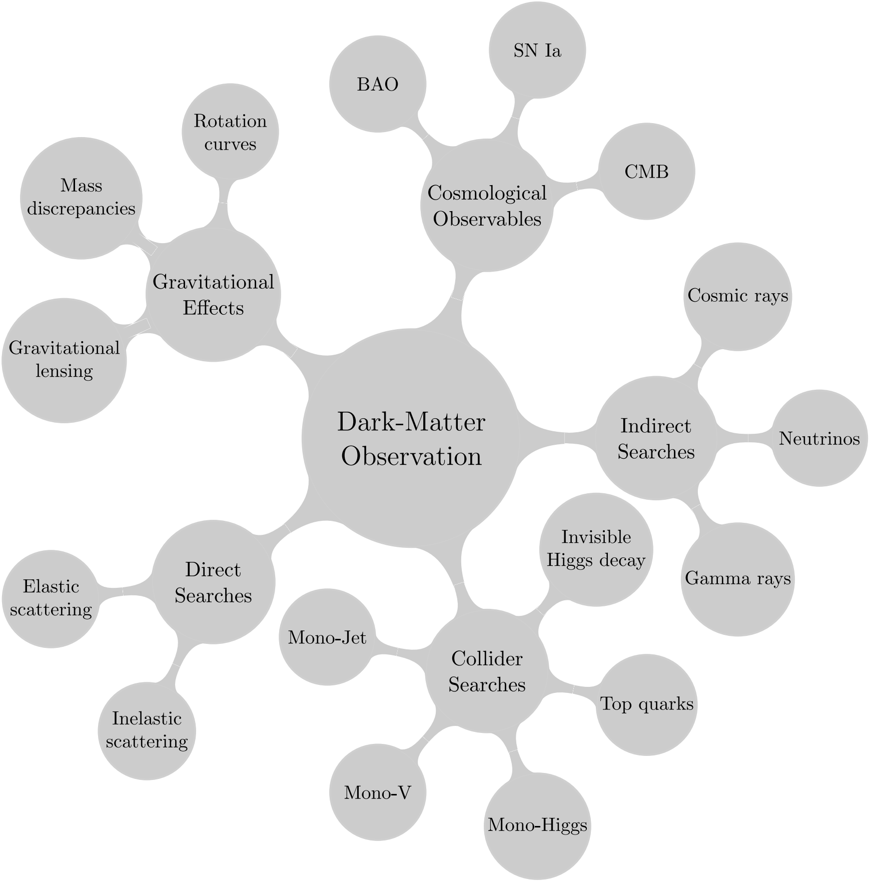

Figure 1. Dark-matter observation: The first level of the diagram shows the possible methods of dark-matter observation (i) gravitational effects and precision measurements of cosmological observables, (ii) direct searches, (iii) indirect searches, and (iv) collider searches. The second level of the diagram shows the various phenomena responsible for each type of observation.

multiple cross-checks regarding the properties of dark matter. Ultimately, the question we are interested in exploring is whether this plurality of methods for probing dark matter can be exploited for the use of robustness arguments from the variability of experiments in order to (1) reduce the number of viable models for the nature of dark matter and (2) increase our confidence about the possible properties of dark matter as captured by specific model scenarios for the particle nature of dark matter. As mentioned in the introduction, robustness arguments cannot be employed for fulfilling aim 1, but they can—and should—be employed for fulfilling aim 2.

To see why, it is helpful to first examine some of the epistemic virtues of the various methods for dark-matter observation in terms of their *informativeness*, their *model sensitivity*, and the *reliability* of the extracted results. The following analysis of these epistemic virtues illustrates how robustness arguments become relevant in dark-matter research and serves as a conceptual foundation for making sense of the fact that different methods of dark-matter observation provide different kinds of information about the properties of dark matter and have varying degrees of sensitivity to the existing viable candidate models for the particle nature of dark matter. This analysis is also necessary for making sense of one of the central claims of this article, namely, that the mere fact that different methods can be used for probing the properties of dark matter is not sufficient for the employment of robustness arguments.

4. Informativeness, model sensitivity, and reliability

A necessary condition for employing robustness arguments from the variability of experiments in order to ensure the reliability of results from the various methods of dark-matter observation is that these methods are actually probing and constraining the same quantities. Siska De Baerdemaeker (2021) highlights this point in her discussion of the implications of methodological pluralism in dark-matter research by noting that "a crucial condition for measurement robustness is that the *same parameter or quantity* is being measured by the different experiments" (140, emphasis added). She then suggests that the common core within the different experiments is provided by the fixed definition of the target of the observation: "the definition of the target system remains fixed under the employment of different methods. It provides, a common core that might underlie multiple methods attempting to probe the same target. Without this agreement on the common core, it is not obvious that methods that detect different phenomena are still probing the same target system and that measurement robustness arguments therefore apply" (140).

Although it is true that all methods just discussed share the common goal of probing the physics of dark matter, and the results they provide do indeed concern the properties of dark matter, as De Baerdemaeker notes, what I aim to show is that the thin definition of dark matter alone does not suffice to ensure that the agreement of results between different methods can be achieved. In order to establish that the results of two different methods are reliable via robustness arguments, we first need to ensure that the extracted information concerns the properties of *the same dark-matter model scenarios and submodels* and relies on the *same assumptions*. However, as will be shown, this is rarely the case in dark-matter research. The concepts of *informativeness* and *model sensitivity* will help us clarify this point.

Informativeness. The informativeness of a method concerns its ability to provide information on a number of different *properties* of dark matter, either by providing specific values for these properties or by constraining the parameter space of a model scenario and/or its submodels. The properties of dark matter for which one can derive information from a particular method could be the cosmic relic density of dark matter, the mass of a dark-matter particle, the self-interaction cross-section, the cross-section between dark matter and baryonic matter, and so on. Assessing the informativeness of a method is crucial for the employment of robustness arguments because such an analysis requires that the involved methods provide information about the same property. However, in practice, the nature and the amount of information about the properties of dark matter that can be inferred from an observational method are typically determined by a number of factors that make the comparison of the information from different methods a much more complicated—if not impossible—process.

In other words, even in cases where two or more different methods are ostensibly providing information about the same property (e.g., constraints on the mass of WIMPs), this information is conditional on a number of factors and assumptions that vary significantly in each experiment. Hence, ensuring that two different methods are probing the same quantity/parameter and provide concordant results requires taking into consideration the effects of these assumptions in the results of the experiment. In the context of dark-matter observation, these factors can be grouped into three different categories: (i) the experimental models of the experiment, (ii) the extrapolating assumptions, and (iii) the model scenarios of dark matter and their submodels.[12]

Experimental models cover a broad category of models referring to every possible modeling activity that facilitates the construction of an experiment and the completion of a measurement process. For instance, in collider experiments, this includes the various physical models and simulations for calculating the interactions of the produced particles with the different parts of the detector, whereas in direct searches, such models would describe the ionization process of the liquid detectors and the interactions of photons with photomultiplier tubes.[13] The interpretation of data from an experiment and the further extrapolation of results strongly depend on the adopted experimental models in the various stages of the experiment because the implementation of different models would give rise to a different set of data. The effects of the adopted experimental models are often implemented in the results in the form of uncertainties, although it is also possible that a number of different results will be derived, based on the selection of a specific combination of models.[14]

Extrapolating assumptions are those assumptions needed for carrying out the required calculations for deriving information about the properties of dark matter *after the acquirement of data from an experiment*. A profound example of a set of extrapolating assumptions comes from direct searches, where the results are necessarily extrapolated on the basis of some standard simplified assumptions, such as the local density ρ_0 of dark matter, an isothermal profile of dark-matter density, and a Maxwell–Boltzmann velocity distribution, incorporated in what is known as the *standard halo model*. In the absence of knowledge about the exact properties of dark matter in the local region, the introduction of these assumptions is essential for carrying out the necessary calculations for the derivation of constraints on various properties of dark matter.

Finally, the information an observational method yields also depends on the model scenario under consideration and its various submodels. For instance, indirect searches are based on the fundamental assumption that dark matter is self-interacting, and its self-annihilation produces standard model particles in the

---

[12] These three categories of factors are not always entirely independent of each other; however, this categorization is helpful to illustrate how the extraction of information from an observational method is conditional on a number of assumptions of different origin.

[13] See Gueguen (2020) and Smeenk and Gallagher (2020) for some nice discussions on the use of simulation in cosmology.

[14] For more detailed discussions on the impact of experimental models on the final results of an experiment, see Antoniou (2021) and Staley (2020).

form of gamma rays, neutrinos, and cosmic rays. Similarly, current collider experiments are only able to provide constraints based on the assumption that dark matter consists of WIMPs that can be produced in high-energy collisions. However, as already mentioned, the model scenario for WIMPs covers a broad class of specific models, and depending on which submodel is taken into consideration, a method can produce more than one set of results, which means that *the derived constraints on a dark-matter candidate from various experiments are highly model dependent in nature*, either in terms of the model scenario or in terms of specific submodels.

Model Sensitivity. The fact that the extrapolation of results from a particular method crucially depends on model-related assumptions implies that the extracted information from a particular method of observation is most of the time *model specific* and associated with either a certain model scenario and/or its submodels. Model sensitivity is the epistemic virtue that concerns the ability of a method to provide model-specific information on a range of different model scenarios of dark matter and their submodels, and it is thus assessed with respect to the number of model scenarios (and their submodels) for which a method can determine their properties and constrain their parameter space. More model-sensitive methods provide information about more dark-matter model scenarios, and vice versa. Hence, although informativeness and model sensitivity are closely related, their difference lies in the fact that the former concerns the number of *properties* of dark matter for which a method yields information (mass, cosmic relic density, cross-sections, etc.), and the latter concerns the number of *model scenarios* and *submodels* of dark matter for which a method is able to provide information (WIMPs, axions, sterile neutrinos, etc.).

A good example of a highly model-sensitive method comes from the indirect searches of dark matter because these experiments are able to provide constraints on a number of different model scenarios, including self-interacting dark matter (SIDM), WIMPs, sterile neutrinos, and models of complex dark matter. On the other hand, the cosmological observation of dark matter via precision measurements of CMB observables is considered to be highly model insensitive because the derived cosmic relic density from this method is insensitive to the various model scenarios of dark matter. In other words, measurements of temperature anisotropies on the CMB provide the current relic density and the stability of dark matter on the cosmological scale, but insofar as it is possible that dark matter is made from more than one component, this information places no model-specific constraints on the relevant model scenarios for the particle nature of dark matter.

The upshot is that when taking into consideration the informativeness of a particular method of dark-matter observation, it is important to highlight the degree to which the constraints imposed by its results are tied to specific model scenarios and/or their submodels. Table 1 illustrates a tentative depiction of the model sensitivity of each method of observation with respect to various model scenarios of dark matter at the present time. A check mark indicates that a method can provide information for at least one parameter of the relevant model, but it should be noted that the situation might well change in the future.[15]

---

[15] For instance, axion-like particles are expected to be searched for in next-generation colliders (Bauer et al., 2019), and there is also a possibility of directly detecting collisionless dark matter via its gravitational effects.

Table 1. Model Sensitivity of the Various Methods of Dark-Matter Observation

	Cosmological	Direct	Indirect	Collider
Collisionless dark matter	✓			
SIDM	✓		✓	
WIMPs	✓	✓	✓	✓
Sterile neutrinos			✓	
Axions		✓	✓	
Hidden/Complex dark matter		✓	✓	✓
Light gravitinos				✓

Reliability. The consideration of the informativeness and model sensitivity of each method provides a good way of evaluating the nature of the information that can be extracted and the various models that can be constrained by each method. What remains to be seen is how these two virtues relate to the reliability of the results from each method because, ultimately, what is of utmost importance for achieving the necessary progress in dark-matter research is whether the extracted results from a method are reliable and can be used to enrich the common core concept of dark matter.

Franklin (1989, ch. 6) nicely summarizes a number of different epistemological strategies that are commonly used in physics experiments in order to justify the reliability of experimental results. As Franklin explains, the implementation of these strategies is often used as a *rational argument to increase the confidence of the scientific community* that the results of an experimental method are reliable and should be used for the construction of further more accurate models. Franklin's discussion includes a number of standard strategies, such as the repetition of an experiment to reproduce the same result and the calibration of instruments based on previous known phenomena; however, the strategy that is most relevant for the purposes of our discussion is what Franklin discusses in terms of the variability of experiments because it is here that robustness arguments become especially useful.

As discussed in section 2, robustness arguments from the variability of experiments are inferences to the best explanation about the agreement of results obtained by two or more experimental methods that are different in nature. The agreement of results from different methods is often used in conjunction with other possible strategies as an additional strong rational argument that increases the scientists' confidence that the common result in these methods is indeed reliable. The main idea is that the best explanation for the fact that two distinct and independent types of experiments provide concordant results is that the material equipment is working properly and the underlying theory and assumptions in these methods are correct; otherwise, as Hacking (1983) aptly noted, "it would be a preposterous coincidence" (as quoted by Franklin 1989, 166) if the two methods produced the same patterns as a result of the same kind of errors. The value of robustness arguments from the variability of experiments thus lies in the fact that they provide further good reason to believe that the common results coming from different methods are reliable.

Section 2 noted that, prima facie, robustness arguments can be used in dark-matter research (1) to reduce the number of viable models for the particle nature of dark matter and (2) to increase the scientists' confidence about the properties of a specific model of dark matter. Ideally, the first aim would be achieved if two or more methods support a specific hypothesis about a property of dark matter, for example, a set of model-independent constraints on the mass of dark-matter particles that excludes certain model scenarios. Such a hypothesis, if supported by the results of two or more methods, would be significantly more robust compared to any competing hypotheses supported by only one method and would thus provide strong grounds to believe that the hypothesis regarding the possible mass of a dark-matter particle is correct. The second and more modest aim requires that model-specific constraints of model scenarios and/or submodels enjoy a higher degree of confidence if they are derived on the basis of two or more methods.

In the next and final section, I shall use the concepts of informativeness, model sensitivity, and reliability to argue that the use of robustness arguments is not a particularly useful strategy for the accomplishment of the first aim; however, it can be used for achieving the second aim, that is, for deriving more reliable model-specific constraints on the viable model scenarios of dark matter and their submodels.

5. The limits and the value of robustness arguments

As already made salient, a necessary condition for the employment of robustness arguments from the variability of experiments is that the results of these methods concern the *same parameter*—that is, they are concordant. Stegenga (2009) argues that although robustness is a valuable epistemic guide in "ideal epistemic circumstances," when it comes to real scientific practice, it faces important limitations. For Stegenga, the main problem with robustness arguments is that, in practice, most of the time, multiple and independent experimental techniques provide results that are inconsistent (i.e., one method suggests x and another method suggests $\neg x$) or incongruent (i.e., one method suggests x and another method suggests y), and hence, it is not clear what kind of epistemic support is provided to the relevant hypotheses. Although this seems to be the case in dark-matter research as well, it will be argued that even in cases where multiple methods ostensibly provide concordant results, the dependence of such results on a number of factors concerning the introduced assumptions in the experiments and the dependence of the experiments on specific dark-matter model scenarios and their submodels makes the task of establishing that the results of different experiments are actually the same extremely difficult. Hence, robustness arguments from the variability of experiments in dark-matter research are limited in scope, not only because the results of multiple methods rarely agree, as Stegenga notes, but also because even in cases of a potential agreement on the surface, establishing the concordance of these results is challenging.

The first and most straightforward complication in establishing that different methods provide concordant results comes from the fact that, as shown in the discussion regarding model sensitivity, compared to the huge variety of model scenarios, there is relatively little overlap between different methods that are able to probe the same model scenarios, let alone the same parameters of these models. Simply put, for a large number of viable candidate model scenarios for dark matter, obtaining results from more than one method is just not possible, and thus robustness arguments cannot be used for the corroboration of hypotheses concerning those models. This is the case, for instance, with sterile neutrinos because, as table 1 shows, they can only be probed via indirect searches.

Moreover, even in the overlapping cases where certain model scenarios can be probed by two or more methods, most of the time, the extracted information concerns fairly small and disjointed areas of the relevant parameter space, making the comparison of results impossible. This is the case, for instance, with WIMPs, which can be better probed in the low-mass regions with colliders, whereas higher masses are better constrained by direct searches. This is a simple example of a trivial case of incongruency, which Stegenga (2009, 654) seems to have in mind when saying that often one method suggests *x* and another method suggests *y*.[16]

However, one of the main aims of this article is to elaborate Stegenga's claim and highlight the subtle fact that even in cases where multiple methods ostensibly provide concordant results (e.g., constraints on the same parameter space for the mass of dark-matter particles), ensuring that these results are indeed in agreement is a very difficult—if not impossible—task. The idea is that incongruency might arise not only from the simple fact that the methods are probing different properties or different areas of the parameter space of a model but also from the fact that the results are (a) model specific for certain model scenarios and (b) conditional on the experimental models of the methods and the extrapolating assumptions. This is a more nuanced form of incongruency that is different than the aforementioned trivial case, in that the incongruence comes from the dependence of the result on different model scenarios and different assumptions, not from the fact that the results are simply about different properties of dark matter.

The main source of this difficulty relates to the informativeness and model sensitivity of each method and concerns the fact that the extracted information from the various different methods is almost always model specific for a particular model scenario and conditional on the experimental models and the extrapolating assumptions of each method. Hence, given that each model scenario comes with its own additional assumptions, constraining, for instance, the mass of WIMPs with colliders (e.g., Goodman et al. 2010) is not the same as constraining the mass of axions with cosmological observations (e.g., Hannestad et al. 2010), despite the fact that both methods provide information about the "mass of a possible dark-matter particle" and the derived constraints on the parameter space may overlap. To ensure the robustness of a new hypothesis H_{DM} concerning a universal property of dark matter that can be included in the common core concept by taking advantage of the variability of methods, such a hypothesis must be model independent with respect to the various model scenarios. This is rarely the case, however. The only model-independent properties of dark matter that are currently available concern the

---

[16] Regarding inconsistency, a well-known example of inconsistent results in dark-matter research comes from the controversial result of the DAMA/LIBRA collaboration (Bernabei et al., 2008) claiming evidence of dark-matter particles in the galactic halo. Subsequent repetitions of the same experiment by different experimental groups (Xenon100, CDMSII, and CoGent) failed to reproduce the same results, thus decreasing the reliability of the initial positive results.

cosmic relic density of dark matter, the upper and lower limits of its particle mass, and the fact that dark matter is cold and dark, which are already encompassed in the thin common core concept of dark matter, assuming standard gravitational laws. The real challenge is to enrich this common core by deriving further properties (e.g., by further constraining the allowed model-independent mass range); however, such properties cannot be derived by using robustness arguments from the variability of experiments due to the model dependence of the results from the available methods on the various model scenarios and their submodels.

This means that robustness arguments can only be employed within a range of different submodels in a specific model scenario, such as the model scenario in which dark matter consists of WIMPs. As we have seen in our discussion of informativeness, in order to extrapolate a meaningful result from an experimental process, numerous assumptions need to be implemented during the construction of the experimental setup and the analysis of data. Ensuring that two different methods are probing the same quantity/parameter of the same model scenario in order to employ robustness arguments thus requires taking into consideration the effects of these assumptions in the results of the experiment.

The following remarks from Goodman et al. (2010) offer a rather illuminating example of how the results of an experiment are conditional to a number of factors. In a paper presenting a set of constraints on properties of dark matter as WIMPs from collider experiments, the authors begin their discussion by stating that the interpretation of the results depends on the nature—and hence the adopted submodel—of the dark-matter particle:

> We consider the cases where the [dark matter] particle is a scalar [boson] or a fermion; if a scalar, it can be real or complex, and if a fermion, it can be Majorana or Dirac. Each of these cases is considered separately. (Goodman et al., 2010, 2)

They then continue by listing the extrapolating assumptions in the experiment in order to yield their results:

> We shall be considering the situation where the WIMP … is the only particle in addition to the standard model fields accessible to colliders … For simplicity, we assume the WIMP is a singlet under the SM gauge groups, and thus possesses no tree-level couplings to the electroweak gauge bosons. We also neglect couplings with Higgs bosons. (Goodman et al., 2010, 2)

After the presentation of their results, Goodman et al. (2010) proceed to conclude that the presented constraints on the strength of interactions of WIMPs with hadrons also depend on the mass of the dark-matter candidate, as well as the coupling preference of dark matter: if dark matter primarily couples to gluons, the constraints from colliders become significantly tighter (8).

These remarks by Goodman et al. (2010) indicate that the various constraints placed on the interactions between WIMPs and hadrons from collider experiments are conditional on a set of introduced assumptions and are, of course, model specific for the model scenario of dark matter as WIMPs. Given that different methods necessarily involve different assumptions, any comparison between these constraints and the constraints obtained by a different method (e.g., from direct searches) must therefore be made by taking into consideration the effects of these assumptions on the extrapolation of the constraints. This is precisely the aim of compatibility and complementarity studies in dark-matter research; however, the severe lack of such studies highlighted by many physicists (e.g., Bauer et al. 2015) indicates the degree of difficulty in achieving this task. The current situation in dark-matter research comprises a vast collection of largely unrelated papers placing *model-specific constraints* on different model scenarios of dark matter without examining the possible concordance of their results with alternative experiments. It is here that robustness arguments from the variability of experiments become valuable and can be used via complementarity studies with the aim of combining the results between different methods in order to improve the reliability and accuracy of the currently available model-specific constraints on dark matter.

A nice example of such work comes from a study by Cerdeno et al. (2016), who present an improvement on dark-matter exclusion limits from direct detection experiments by incorporating the results from indirect searches based on gamma rays in the Milky Way halo. Interestingly, the authors begin the abstract of their article by stating, “When comparing constraints on the Weakly Interacting Massive Particle (WIMP) properties from direct and indirect detection experiments it is crucial that the assumptions made about the dark matter (DM) distribution are realistic and *consistent*” (00, emphasis added). They then proceed to calculate a consistent and improved model-specific exclusion limit on the (WIMP) dark-matter nucleon-scattering cross-section by taking into consideration the introduced assumptions about the dark-matter contribution in both direct and indirect searches. Such studies nicely illustrate how robustness arguments from the variability of experiments can be used in dark-matter research in the context of complementarity and compatibility studies to improve the reliability of the constraining limits in the parameter space of certain submodels of a given model scenario—in this case, within the model scenario of dark matter as WIMPs.

There are, however, two important caveats. The first is that such results are—as noted—always model specific because they are limited to the various submodels of certain model scenarios, and as such, they leave the essential question about which model (or models) best captures the actual nature of dark matter untouched. The second caveat is that such studies are only available for the submodels of those model scenarios that can be probed by more than one method, precisely because the results from different methods are dependent on the same model scenario and thus share a common set of assumptions that allows them to be compared. Hence, robustness arguments from the variability of experiments in the form of compatibility studies are of limited use to model scenarios that can be currently probed by only one method, such as sterile neutrinos and light gravitinos. Nevertheless, robustness arguments from the variability of experiments can play a significant role in improving the reliability of results concerning the allowed parameter space of the various competing models, insofar as the different additional assumptions on which the results of each method are dependent are taken into consideration during the comparison of concordant results.

6. Conclusions

In summary, the aim of this article was to examine the extent to which robustness arguments from the variability of experiments can be used within dark-matter research in order to overcome various methodological and theoretical challenges by exploring the concepts of informativeness, model sensitivity, and reliability as useful epistemic virtues. Apparently, the most important challenge in dark-matter research is to reduce the number of viable models of dark matter by enriching the minimum set of constraints in the thin common core concept in a way that eliminates certain candidate models. The first conclusion is that robustness arguments are of limited use in achieving this aim because most of the time, the extracted information from the various methods of dark-matter observation is model specific. The extraction of model-independent constraints from a particular method—let alone from two or more methods for the needs of robustness arguments—remains one of the most important challenges in dark-matter research.

The second conclusion is that robustness arguments from the variability of experiments can—and should—be used for constraining the parameter space of certain model scenarios via compatibility and complementarity studies. This can be achieved by extracting exclusion limits for the properties of those model scenarios that can be probed by more than one method by taking into consideration the different assumptions introduced in each method and making sure that they are consistent. The upshot is that although robustness arguments from the variability of experiments cannot be employed for the elimination of possible model scenarios, they are still a valuable tool for increasing the confidence of the scientific community about the reliability of the current methods of dark-matter observation and the allowed parameter space of certain models, thus contributing to the achievement of a more modest, but nonetheless important, aim.

Acknowledgments. I am indebted to Karim Thébault for our insightful discussions on dark matter and for a set of constructive comments on earlier versions of this work. Many thanks also to Niels Martens and the audience of an online workshop on the epistemology of dark matter at the University of Bonn, for their constructive questions and their feedback on some of the ideas contained in this work. This research was supported by an Arts and Humanities Research Council Ph.D. Studentship, awarded through the South, West and Wales Doctoral Training Partnership.

Competing interests. The author declares that there is no competing interest.

References

Antoniou, Antonis. 2021. “What Is a Data Model?” *European Journal for Philosophy of Science* 11 (4):1–33.

Basso, Alessandra. 2017. “The Appeal to Robustness in Measurement Practice.” *Studies in History and Philosophy of Science Part A* 65:57–66.

Bauer, Daniel, James Buckley, Matthew Cahill-Rowley, Randel Cotta, Alex Drlica-Wagner, Jonathan L. Feng, Stefan Funk, JoAnne Hewett, Dan Hooper, and Ahmed Ismail. 2015. “Dark Matter in the Coming Decade: Complementary Paths to Discovery and Beyond.” *Physics of the Dark Universe* 7:16–23.

Bauer, Martin, Mathias Heiles, Matthias Neubert, and Andrea Thamm. 2019. “Axion-Like Particles at Future Colliders.” *European Physical Journal C* 79 (1):74.

Bernabei, R., P. Belli, F. Cappella, R. Cerulli, C. J. Dai, A. D’Angelo, H. L. He, A. Incicchitti, H. H. Kuang, and J. M. Ma. 2008. “First Results from DAMA/LIBRA and the Combined Results with DAMA/NaI.” *European Physical Journal C* 56 (3):333–55.

Bertone, Gianfranco, Dan Hooper, and Joseph Silk. 2005. "Particle Dark Matter: Evidence, Candidates and Constraints." *Physics Reports* 405 (5–6):279–390.

Boyarsky, Alexey, Marco Drewes, Thierry Lasserre, Susanne Mertens, and Oleg Ruchayskiy. 2019. "Sterile Neutrino Dark Matter." *Progress in Particle and Nuclear Physics* 104:1–45.

Cerdeno, David G., Mattia Fornasa, Anne M. Green, and Miguel Peiró. 2016. "How to Calculate Dark Matter Direct Detection Exclusion Limits That Are Consistent with Gamma Rays from Annihilation in the Milky Way Halo." *Physical Review D* 94 (4):43516.

De Baerdemaeker, Siska. 2021. "Method-Driven Experiments and the Search for Dark Matter." *Philosophy of Science* 88 (1):124–44.

De Baerdemaeker, Siska and Nora Mills Boyd. 2020. "Jump Ship, Shift Gears, or Just Keep on Chugging: Assessing the Responses to Tensions between Theory and Evidence in Contemporary Cosmology." *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 72:205–16.

de Swart, Jaco. 2020. "Closing in on the Cosmos: Cosmology's Rebirth and the Rise of the Dark Matter Problem." In *The Renaissance of General Relativity in Context*, edited by Alexander Blum, Roberto Lalli, and Jürgen Renn, 257–84. Cham, Switzerland: Springer.

Duffy, Leanne D., and Karl Van Bibber. 2009. "Axions as Dark Matter Particles." *New Journal of Physics* 11 (10):105008.

Feng, Jonathan L. 2010. "Dark Matter Candidates from Particle Physics and Methods of Detection." *Annual Review of Astronomy and Astrophysics* 48:495–545.

Franklin, Allan. 1989. *The Neglect of Experiment*. Cambridge: Cambridge University Press.

Gaskins, Jennifer M. 2016. "A Review of Indirect Searches for Particle Dark Matter." *Contemporary Physics* 57 (4):496–525.

Gluscevic, Vera, Yacine Ali-Haimoud, Keith Bechtol, Kimberly K. Boddy, Céline Bœhm, Jens Chluba, Francis-Yan Cyr-Racine, Cora Dvorkin, Daniel Grin, and Julien Lesgourgues. 2019. "Cosmological Probes of Dark Matter Interactions: The Next Decade." arXiv preprint.

Goodman, Jessica, Masahiro Ibe, Arvind Rajaraman, William Shepherd, Tim M. P. Tait, and Hai-Bo Yu. 2010. "Constraints on Dark Matter from Colliders." *Physical Review D* 82 (11):116010.

Gueguen, Marie. 2020. "On Robustness in Cosmological Simulations." *Philosophy of Science* 87 (5):1197–208.

Hacking, Ian. 1983. *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge: Cambridge University Press.

Hannestad, Steen, Alessandro Mirizzi, Georg G. Raffelt, and Yvonne Y. Y. Wong. 2010. "Neutrino and Axion Hot Dark Matter Bounds after WMAP-7." *Journal of Cosmology and Astroparticle Physics* 2010 (8).

Jacquart, Melissa. 2021a. "Dark Matter and Dark Energy." In *Routledge Companion to the Philosophy of Physics*, edited by Eleanor Knox and Alastair Wilson, 731–43. New York: Routledge.

Jacquart, Melissa. 2021b. "ΛCDM and MOND: A Debate about Models or Theory?" *Studies in History and Philosophy of Science Part A* 89:226–34.

Kahlhoefer, Felix. 2017. "Review of LHC Dark Matter Searches." *International Journal of Modern Physics A* 32 (13):1730006.

Kosso, Peter. 2013. "Evidence of Dark Matter, and the Interpretive Role of General Relativity." *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 44 (2):143–47.

Lisciandra, Chiara. 2017. "Robustness Analysis and Tractability in Modeling." *European Journal for Philosophy of Science* 7 (1):79–95.

Lloyd, Elisabeth A. 2015. "Model Robustness as a Confirmatory Virtue: The Case of Climate Science." *Studies in History and Philosophy of Science Part A* 49:58–68.

Luković, Vladimir, Paolo Cabella, and Nicola Vittorio. 2014. "Dark Matter in Cosmology." *International Journal of Modern Physics A* 29 (19):1443001.

Martens, Niels C. M. 2022. "Dark Matter Realism." *Foundations of Physics* 52 (1):1–19.

Martens, Niels C. M., and Dennis Lehmkuhl. 2020. "Dark Matter = Modified Gravity? Scrutinising the Spacetime–Matter Distinction through the Modified Gravity/Dark Matter Lens." *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 72:237–50.

Massey, Richard, Thomas Kitching, and Johan Richard. 2010. "The Dark Matter of Gravitational Lensing." *Reports on Progress in Physics* 73 (8):86901.

Massimi, Michela. 2018. "Three Problems about Multi-Scale Modelling in Cosmology." *Studies in History and Philosophy of Modern Physics* 64:26–38.

Parker, Wendy S. 2011. "When Climate Models Agree: The Significance of Robust Model Predictions." *Philosophy of Science* 78 (4):579–600.

Roszkowski, Leszek, Enrico Maria Sessolo, and Sebastian Trojanowski. 2018. "WIMP Dark Matter Candidates and Searches—Current Status and Future Prospects." *Reports on Progress in Physics* 81 (6):66201.

Schumann, Marc. 2019. "Direct Detection of WIMP Dark Matter: Concepts and Status." *Journal of Physics G: Nuclear and Particle Physics* 46 (10):103003.

Schupbach, Jonah N. 2018. "Robustness Analysis as Explanatory Reasoning." *British Journal for the Philosophy of Science* 69 (1):275–300.

Shapere, Dudley. 1982. "The Concept of Observation in Science and Philosophy." *Philosophy of Science* 49 (4):485–525.

Smeenk, Chris. 2020. "Some Reflections on the Structure of Cosmological Knowledge." *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 71:220–31.

Smeenk, Chris, and Sarah C. Gallagher. 2020. "Validating the Universe in a Box." *Philosophy of Science* 87 (5):1221–33.

Staley, Kent W. 2020. "Securing the Empirical Value of Measurement Results." *British Journal for the Philosophy of Science* 71 (1):87–113.

Stegenga, Jacob. 2009. "Robustness, Discordance, and Relevance." *Philosophy of Science* 76 (5):650–61.

Vanderburgh, William L. 2014. "On the Interpretive Role of Theories of Gravity and 'Ugly' Solutions to the Total Evidence for Dark Matter." *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 47:62–67.

Weisberg, Michael. 2006. "Robustness Analysis." *Philosophy of Science* 73 (5):730–42.

Weisberg, Michael, Melissa Jacquart, Barry Madore, and Marja Seidel. 2018. "The Dark Galaxy Hypothesis." *Philosophy of Science* 85 (5):1204–15.

Woodward, Jim. 2006. "Some Varieties of Robustness." *Journal of Economic Methodology* 13 (2):219–40.

Cite this article: Antoniou, Antonis. 2023. "Robustness and Dark-Matter Observation." *Philosophy of Science* 90 (3):629–647. https://doi.org/10.1017/psa.2023.50