Scientific Understanding in the Aharonov-Bohm Effect

By Elay Shech

Department of Philosophy
6080 Haley Center
Auburn University, AL 36849-3715

ABSTRACT

By appealing to resources found in the scientific understanding literature, I identify in what senses idealizations afford understanding in the context of the (magnetic) Aharonov-Bohm effect. Three types of concepts of understanding are discussed: understanding-what, which has to do with understanding a phenomenon; understanding-with, which has to do with understanding a scientific theory; and understanding-why some phenomenon occurs. Consequently, I outline an account of understanding-with that is suggested by the historical controversy surrounding the Aharonov-Bohm effect.

Forthcoming in *Theoria*

Please cite published version

Acknowledgements: Thank you to John Earman, Guido Bacciagaluppi, Guy Hetzroni, Kareem Khalifa, Arnon Levy, Michael Cuffaro, participants at “Understanding, Causation, and Representation: A Workshop in Philosophy of Science” (at the Sidney M. Edelstein Center at the Hebrew University, 27 December 2017), and participants at “Representation, Idealization, and Explanation in Science” (at Auburn University, 5-6, April, 2019) for helpful discussion and comments on earlier drafts of this paper.

Scientific Understanding in the Aharonov-Bohm Effect

1. Introduction

Speaking to the Royal Society of London in 1803, Thomas Young described his famous double-slit experiment in which a beam of homogenous light falls on a screen with two slits and leads to an interference pattern with dark and light fringes.[1] Young's work marked an important shift from the corpuscular theory of light to the wave theory of light. In the 20th century, it became clear that the wave nature of light extends to matter more generally, as it is possible to produce interference patterns with particles like electrons. Namely, when a coherent beam of electrons is shot at a double-slit screen, an interference pattern of light and dark fringes emerges (Fig. 1; Left), which arises from the buildup of single electrons (Fig. 1; Right).[2] *Physics World's* article on the top ten most beautiful experiments ranked the "double slit experiment with electrons" highest as it "exemplifies the wave–particle duality" of matter and of "quantum physics itself" (Crease 2002, 20).

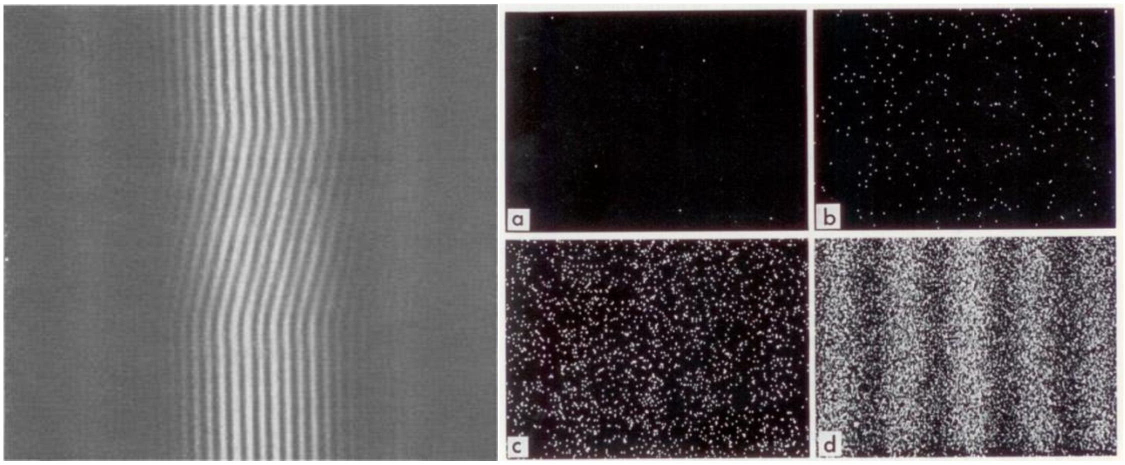

Fig. 1 (Left) An example for an interference pattern from a double-slit experiment (from Möllenstedt and Bayh 1962, 304). (Right) Single-electron build-up of (biprism) interference pattern (from Tonomura 1999, 15). (a) 8 electrons, (b) 270 electrons, (c) 2000 electrons, and (d) 60,000 electrons.

It is interesting then that there is a variant of the double-slit experiment which is, arguably, also fundamental for our understanding of the quantum realm. Specifically,

---

[1] In fact, the experiment that Thomas Young described didn't have double-slits and it is only in 1807, (in his *Course Lectures on Natural Philosophy*) that there is explicit mention of multiple slits. Still, the discovery of interference in light was solidified in his experiments and publications in the period of 1801-1803, culminating with his theory of the interferometer in 1817.

[2] Interestingly, although Clinton Davisson and Sir George Paget Thompson (the son of Sir J. J. Thomson) received the 1937 Noble prize in Physics for "their discovery of the interference phenomena arising when crystals are exposed to electronic beams" (Pleijel 1937)—thereby confirming the wave nature of electrons—the actual double-slit experiment with electrons was not performed until 1960 in Claus Jönsson's experiment (Jönsson 1961).

imagine introducing a magnetic field into the double-slit experimental scenario. As the electrons navigate their course through the apparatus, the Lorentz force will act on them (in this case, due to the presence of charged particles moving through a magnetic field) and the interference pattern emerging on the detector screen will be shifted. Such a shift in pattern can be understood within the context of classical physics as the product of the interaction between the magnetic field and the electron beam.[3] But now add an embellishment: consider what would happen if the magnetic field was *completely* shielded from the electron beam. One can accomplish such a feat by making use of idealizations: introduce a cylindrical solenoid that is infinitely long and absolutely impenetrable to external electrons, and then turn it "on" so that it produces a shielded magnetic field. What happens to the interference pattern in this situation? Classically, we would not expect any special effect to manifest since there is no region in space in which the electron beam and magnetic field can causally interact with each other.

What is remarkable about Yakir Aharonov and David Bohm's 1959 celebrated paper, "Significance of electromagnetic potentials in the quantum theory," is that they were able to show that quantum mechanics predicts a shift in interference pattern in the above stated idealized scenario, and that such an effect can ostensibly be confirmed in the laboratory. Aharonov and Bohm (1959) became immensely influential (e.g., *Physical Review* counts 4855 citations, and Google Scholar counts 8807 citations), but also sparked a heated thirty-year controversy in the physics literature. The debate concerned both the nature and reality of (what has become to be known as) the (magnetic) "AB effect," as well as various foundational issues. Philosophers of physics have been attentive to the AB effect (e.g., Batterman (2003), Belot (1998), Maudlin (1998), Mattingly (2006) Healey (1997, 1999, 2007), Wallace (2014), Daugherty (2021)), and some recent contributions (viz., Shech (2018), Earman (2019)) have focused specifically on the role that idealizations play, suggesting that they provide genuine *understanding*. For instance, Earman (2019, 1991) holds that the point of studying the idealized AB effect "is to understand the foundations of quantum mechanics and how its predictions depart from those of classical mechanics." That said, there has been no connection drawn between the claims of how idealizations afford understanding in the AB effect, on the one hand, and the important concepts and distinctions found in burgeoning literature on scientific understanding (e.g., De Regt *et al*. (2009), De Regt (2017), Grimm *et at*. (2017), Khalifa (2017), Khalifa *et al*. (Forthcoming), Potochnik (2017)), on the other hand. This is so even in light of the fact that the AB effect is a significant and fundamental effect of one of our best scientific theories, viz., quantum mechanics.

The aim of this paper is to attend to this lacuna and take the first steps in filling said gap by pursuing two interrelated goals. First, on the philosophy of physics side, I will flesh out the manner by which idealizations in the AB effect afford understanding by appealing to key concepts found in the epistemological and philosophy of science literature on

---

[3] Although, the presence of an interference pattern on the detector screen due to electrons would certainly be a surprise in a classical context.

understanding. Second, on the philosophy of understanding side, I hope to make a modest contribution by highlighting some ostensible take-home messages suggested by the AB effect historical case study with special attention paid to two notions: understanding-with, viz., (objectual) understanding of a scientific theory, and understanding-what, viz., (objectual) understanding of a phenomenon. The basic idea can be summed up as follows: Scientific understanding ought to track scientific practice, and the AB effect case study shows that there are possible (but non-actual) worlds according to quantum mechanics, which are most naturally construed as idealizations, and are essential for understanding both quantum mechanics and what the AB effect actually is in the first place.[4]

The structure of the paper is as follows. Section 2 will set the stage, so to speak, by both introducing some important concepts and distinctions found in the scientific understanding literature and situating my own use of terms like "understanding-with" and "understanding-what." In Section 3, I describe features of the AB effect and aspects of the historical controversy needed to substantiate my claims. Additional technical details are left for footnotes and an Appendix. Section 4 will discuss the concept of understanding-with, while Section 5 outlines an account of understanding-with that is suggested by the AB effect case study. Section 6 discusses the concept of understanding-what. In Section 7, I briefly summarize claims and draw implications for understanding-why the AB effect occurs. Section 8 considers and responds to various objections, while Section 9 ends the paper with a short conclusion.

Before starting in earnest, it would be worthwhile to note two caveats. First, an "idealization" is usually characterized as a distortion or falsehood, and it is contrasted with an "abstraction," in which no falsehood is asserted since details are abstracted away. Both notions are further distinguished from an "approximation," which tends to concern a purely formal matter (such as calculation for some purpose) and is thus not usually taken to present interpretive issues. An idealized or abstracted model, system, phenomenon, or scenario is typically referred to as a "fiction," to draw attention to the fact that, strictly speaking, no such actual object exists; or as something that is "abstract" to highlight the contrast with concrete, physical objects. A world that contains idealizations can also be referred to as an idealization of sorts. For my purposes, unless noted otherwise, I will take a broad view of the notion of idealization to include all of the above and accept any over-generation (or under-generation) that results.[5]

---

[4] Cf. Belot (1998) who discusses how the AB effect and quantum consideration enhance and elucidate our understanding of classical electromagnetism, as well as shape our interpretation: "When we discover that holonomies are preferable to magnetic fields for interpreting [classical] electromagnetism, … we do discover that one interpretation provides a more fruitful way of thinking about a certain range of phenomena than the other—both provide a satisfying picture of electromagnetism, but only one of them provides helpful hints about the quantum realm" (553-554).

[5] For instance, Norton (2012, 209) distinguishes between an "approximation," which is an "inexact description of a target system," and an "idealization," which is "a real or fictitious system, distinct from the target system, some of whose properties provide an inexact description of some aspects of the target system." In Norton's terminology then, I will be concerned with idealizations, viz., specific

Second, I will use "practically essential," or "essential" for short, as a kind of term of art, identifying a midway position between the concepts of "necessity" and "sufficiency," as these concepts are too coarse grained for my purposes. By "essential" I mean something along the lines of strong pragmatic necessity, as opposed to a weak necessity arising from issues of tractability, ease of calculation, simplification, and convenience. Consider a motivating example: If one were to claim that "the thermodynamic limit (TDL) is *necessary* for the explanation of phase transitions (PT)," then one is claiming that on any reasonable account of explanation and characterization of PT, and given any acceptable explanation, it is *impossible* to explain PT without appealing to the TDL. This is an exceedingly strong claim that would necessitate something like an in-principle argument to the effect that current and future physics will not be able to explain PT without the TDL, or else a subtle explication of how the TDL is inextricably involved in the very concept of PT. In contrast, if one claims that "the TDL is *essential* for the explanation of PT," then one is claiming that, as a matter of fact, all of our most successfully explanations of PT appeal the TDL, that the TDL is also carrying the explanatory load in these explanations (instead of simply affording tractability or convenience), and it isn't readily clear how one would explain PT without appealing to the TDL. In other words, there is no nearby successful explanation of PT without the TDL in the offing. This is the manner by which I'll use the term "essential."[6]

2. Setting the stage: Scientific Understanding

In this section, I present two families of distinctions found in the understanding literature and explain my use of "understanding-with" and "understanding-what" in relation to said distinctions, thereby outlining the scope of the present project. I follow the norm of using "p" and "q" (or "P") to symbolize either a phenomenon or scenario/state of affairs; or else a proposition describing a phenomenon, scenario, etc. (wherein appropriate use can be extrapolated from context).

To start, Strevens (2013) distinguishes between "understanding-why," "understanding-with," and "understanding-that." Understanding-why concerns having understanding of why some phenomenon occurs, e.g., understanding why p may be constituted by knowing the explanation for p (Strevens 2013, Khalifa 2017). Understanding-with concerns understanding a scientific theory. For instance, on Strevens' (2013, 513) account "understanding with involves mastering a scientific explanation: to understand a theory in this new sense is to be able to use the theory to explain a range of phenomena." Understanding-that, e.g., understanding that the kettle is boiling, "is the fundamental relation between mind and world, in virtue of which the mind has whatever familiarity it does with the way the world is" (511).[7] For the purposes of this paper, I will not treat understanding-that on such a fundamental level. Instead, I adopt a deflationary

---

possible but non-actual worlds (according to quantum mechanics) that are most naturally construed as idealizations (since they contain, e.g., infinitely long solenoids).

[6] However, note that I used PT solely as a clarificatory example; I don't take a stance on the issue here.

[7] Baumberger *et al*. (2017, 5-6) similarly talks about "symbolic understanding," which has to do with "what is meant by linguistic expressions or by other symbols."

view that holds a close relation between understanding-that p and knowing-that p (thereby leaving the deeper foundational study for others).

Second, it is has become common to distinguish between practical understanding, explanatory understanding, and objectual understanding (e.g., Struart (2018) and Baumber *et al*. (2017)). Practical understanding is a kind of know-how or tacit knowledge such as playing the guitar (Khalifa 2013, 123; Stuart 2018, 529). In the scientific context, such understanding is probably most associated with problem solving and applying scientific theory—activities for which idealizations are indispensable. Nevertheless, since the heuristic or pedagogical role of idealization is not my concern, I leave aside this sense of understanding.

Explanatory understanding concerns having an explanation for why p (De Regt 2009, 25; Khalifa 2011, 108; Stuart 2018). Explanatory understanding and understanding-why overlap for accounts of understanding-why that stress the centrality of scientific explanation. For an example of two well-received views consider, first, De Regt's (2017) account, which is articulated in his Criterion of Understanding Phenomena (CUP): "A phenomenon P is understood scientifically if and only if there is an explanation of P that is based on an intelligible theory T and conforms to the basic epistemic values of empirical adequacy and internal consistency" (92). Second, Khalifa (2017) defends the view that understanding is constituted by knowledge of relevant explanatory information. His Explanation-Knowledge-Science (EKS) model of understanding holds that "q (correctly) explains why p if and only if: p is (approximately) true; q makes a difference to p; q satisfies your ontological requirements (so long as they are reasonable); and q satisfies the appropriate local constraints" (Khalifa 2017, 7).[8] For my purposes, however, neither explanatory understanding nor understanding-why will be of *direct* focus. I will discuss these notions and related accounts *indirectly*, as they are related to other concepts of interest. Roughly, the reason is that I do not wish to enter into the highly contested debate of what (causally-mechanistically) explains the AB effect since this would draw attention from the main themes that I aim to emphasize.

Instead, I will be concerned with the notion of "objectual understanding," which has to do with understanding a thing or a domain of things, broadly construed, so that one can grasp and express the relations among events, objects, domains, etc., in an appropriate manner. There is some disagreement as to the nature of objectual understanding and its object. Some folks take objectual understanding to concern understanding a topic or subject matter (e.g., Brun and Baumberger (2017), Carter and Gordon (2016), Elgin (2017)). For example, Catherine Elgin notes:

> An understanding […] is an epistemic commitment to a comprehensive,

---

[8] It is noteworthy, however, that although the notion of explanation is central to the CUP account, De Regt does not reduce understanding to knowledge as Khalifa's account ostensibly does since De Regt holds that understanding involves skills. For an interesting recent exchange regarding their respective accounts see De Regt (Forthcoming a, Forthcoming b) and Khalifa (Forthcoming a, Forthcoming b).

> systematically connected body of information that is grounded in fact, is duly responsive to evidence, and enables non-trivial inference, argument and perhaps action pertaining to the phenomena the information is about. … The immediate object of such understanding is not a proposition or fact, but a topic or subject matter (Elgin 2017, 82).

Still, others take objectual understanding to mean understanding a theory (e.g., Baumberger (2019), De Regt (2017), Newman (2017)). For instance, Newman's (2017, 582) "inferential model of scientific understanding" (IMU) holds that: [Subject] S understands scientific theory T if and only if S can reliably use principles, Pn, constitutive of T to make goal-conducive inferences … that reliably results in solutions to qualitative problems relevant to that theory." This notion of objectual understanding overlaps with the notion understanding-with that I take Strevens (2013) to have in mind and that I will be working with myself. Namely, I will argue that the idealizations in the AB effect are essential for an objectual understanding of quantum mechanics (Section 4) and show how the case study suggests an outline of a theory of understanding-with (Section 5).

Last, some take objectual understanding to mean understanding a phenomenon or thing (while resisting reduction to explanatory understanding or understanding-why) (Dellsén 2020, Kelp 2015). For instance, Baumberger (2011, 77-s) notes:

> Understanding global warming involves, for instance, understanding what effects (on natural and social systems) it will have, how it is linked to human activities (such as burning fossil fuels and deforestation) and related phenomena (such as the destruction of stratospheric ozone and global dimming), how far greenhouse gas emissions and, as a result, temperatures are likely to rise in the future, and how the changes will vary over the globe. (Baumberger 2011, 77-S)

Interestingly, the notions of objectual understanding of a theory (viz., understanding-with) and objectual understanding of a phenomenon can come apart. One can, for example, have an understanding of Newton's gravitational theory, and yet not really understand gravity because Newton's theory (where gravity is a force) has been superseded by Einstein's theory (where gravity is the curvature of spacetime). However, as an anonymous reviewer notes, some may worry that said notions are interdependent because of, ostensibly, the (at least minimal) theory-ladenness of observation and confirmation. Relatedly, it is crucial to notice that having objectual understanding of certain types of phenomena is strongly entangled with the manner by which the same phenomena are characterized by their governing theory (quantum entangled being a case in point). Since this is the type of objectual understanding of phenomenon that I'm interested in, I'll introduce the term "understanding-what"—as in, understanding what a phenomenon is *in the first place*—to signify this point.

In particular, some phenomena studied by science are such that understanding-what p is straightforwardly related to understanding-that (or knowing-that) p is happening or has occurred. For example, say a scientist wants to study balls rolling down an inclined plane, or a boiling kettle. Understanding-what such phenomena are *in the first place* does not seem to necessitate anything over or above the abilities manifested in virtue of understanding-that (or knowing-that) "balls are rolling down an inclined place" or "a kettle is boil." Specifically, the ability to (say) recognize the relevant observable phenomena may be sufficient. Importantly, one need not appeal to complicated concepts arising in scientific theories.

In contrast, understanding-what PT are does seem to necessitate something over and above, say, understand-that (or knowing-that) a kettle is boiling. In fact, one needs more than understand-that (or knowing-that) iron has magnetized, graphite has spontaneously converted into a diamond, or that a semiconductor transitioned into a superconductor—all examples of PT. It is exceedingly difficult to characterize PT in a theory independent manner, and it isn't clear that understanding-what PT are, *in the first place*, can be achieved without knowing concepts such the partition function, (Helmholtz or Gibbs) free energy, critical exponents, symmetry breaking, etc. After all, on the face of it, a boiling kettle and a superconductor don't seem to have much in common. In any case, whatever one thinks of PT, I will suggest below that the AB effect is an example of the type of phenomenon that is difficult to capture in a theory-independent manner. Furthermore, I will argue that idealizations are essential for understanding-what the AB effect is in the first place (Section 6).

In sum, my focus will be on two notions. First, by "understanding-with" I mean objectual understanding of a theory like quantum mechanics. Second, by "understanding-what" I mean objectual understanding of what a phenomenon like the AB effect is in the first place. I set aside practical understating, objectual understanding of a topic/subject matter, explanatory understanding, and understanding-why p, although I will re-connect with the last two notions in sections 5 and 7-8. Still, one may worry that appealing to the AB effects needlessly overcomplicates my case since all insights regarding idealizations and understanding may be attenable using a simple example of idealizations such as balls rolling down an inclined place (as discussed above). This is not the case. For one, the frictionless plane example is not a case from the history of science where appeals to idealizations were made in order to shed light understanding-with as suggested below. Nor is it an example of a phenomenon for which appeals to theory and idealizations are essential for understanding-what the phenomenon is in the first place.

3. The (magnetic) AB Effect(s): Abstract Vs. Concrete; Narrow vs. Broad. [9]

In a double-slit experiment (as in Fig. 2), a coherent beam of electrons (symbolized by ψ) is shot at a double-slit screen and split into two components (symbolized by ψ_1 and ψ_2), each passing on opposite sides of a cylindrical solenoid, and then made to recombine at the detector-screen where an interference patterns arises (as in Fig. 1). The solenoid is turned "on" and the interference pattern shifts by some amount Δx that systematically depends on the magnetic flux (which is essentially the amount of magnetic field passing through a surface area). Roughly, following Earman (2019), this shift in pattern due to the shielded flux is what we call the AB effect narrowly construed or *narrow AB effect* for short.[10] Three

---

[9] My discussion here and in the Appendix follows Earman (2019) and Shech (2018) closely. Ehrenberg and Siday (1949) are usually credited with first noting the effect and Chambers (1960) with being the first experimental confirmation. Tonomura *et al*. (1982, 1986) are considered the first definitive experimental confirmation of the effect.

[10] It should be stressed that the AB effect discussed here is known as the magnetic AB effect. There is an analogue electric AB effect, implied by Lorentz covariance, which arises from electric (instead of magnetic) fields, but I will concentrate on the magnetic AB effect in this paper (and so by "AB effect" I mean the magnetic AB effect). In addition, foreshadowing what is to come, the narrow AB effect is to be contrasted with the broad AB effect, which concerns observable effects predicted given different boundary conditions.

conspicuous idealizations are appealed to in this scenario (which is based on a variant of Aharonov & Bohm (1959)):

- I1. The cylindrical solenoid is assumed to be infinitely long, so that no magnetic field spills outside of the solenoid once it is already "on."
- I2. When the solenoid is turned "on" the magnetic field generated is completely contained within the solenoid, so that no magnetic field spills outside of the solenoid as it is starting up.
- I3. The solenoid is absolutely impenetrable so that the probability of finding an external electron in the region inside the solenoid is zero.

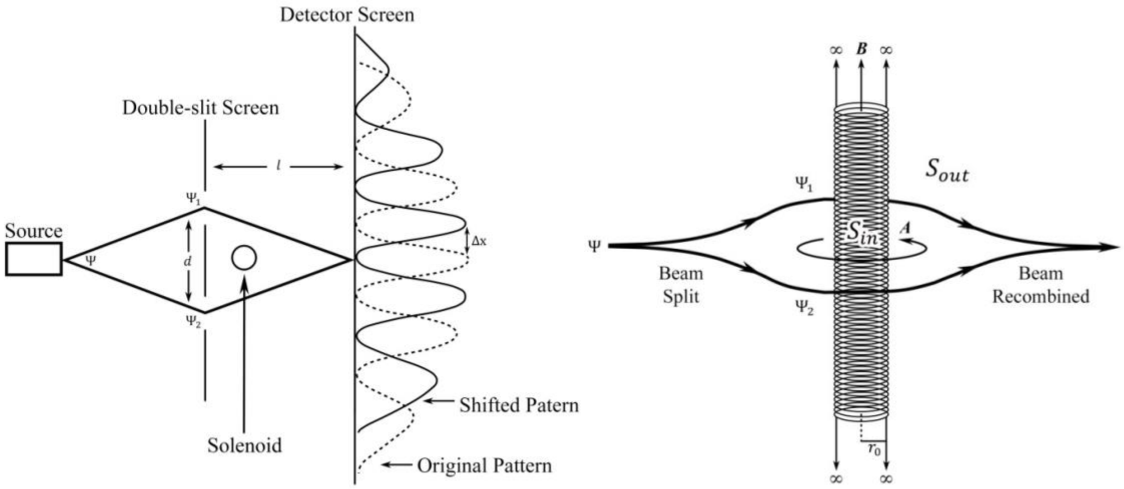

Fig. 2 An illustration of the (magnetic) AB effect.

These three idealizations (I1-I3) collectively ensure that the magnetic field and flux generated by the solenoid is *completely* shielded from the electron beam and vice versa. Making use of such idealizations, as well as the Schrödinger equation corresponding to said setup (which one can think of as a basic law of quantum mechanics describing the dynamical evolution of systems), Aharnov & Bohm (1959) were able to derive the narrow AB effect.[11] However, for practical reasons, it is not possible to build an experimental setup of the sort since, e.g., "infinitely repulsive barriers do not really exist" (Magni and Valz-Gris 1995, 179–180). Accordingly, I distinguish between the fictional, idealized, and theoretical effect, what I will call the *abstract AB effect*—i.e., the AB effect as it is conventionally defined—on the one hand, and its physical counterpart, the *concrete AB effect* that has been allegedly confirmed in the laboratory (since, it is claimed, experiments have come close enough to the ideal scenario).

---

[11] See Appendix. Aharonov and Bohm's (1959) derivation (which further idealized the solenoid radius by shrinking it to zero without affecting the value of the flux) also made use of a conjecture dubbed the Aharonov-Bohm *Ansatz* (Ballesteros and Weder 2009, 2011; Shech 2018, Earman 2019).

Our next distinction concerns a broader conception of the AB effect and necessitates further examination of the formal implementation of idealizations I1-I3. Standardly, idealizations are means to *simplify* problem-situations that would be difficult or intractable without simplification. However, I1-I3 are *not* just simplifying idealizations. Instead, they also complicate matters because it turns out that there are many ways to formally implement I1-I3. Specifically, theorists are required to make a choice about the boundary conditions that obtain at the infinite solenoid boundary. We can think about these conditions as describing the manner by which the electron beam interacts with the solenoid border. In the non-idealized case, there is a *unique* dynamical evolution corresponding to the situation at hand and issues having to do with the formal implementation of I1-I3 need not arise.[12] But in the idealized case there are different formal implementations of I1-I3, corresponding to different boundary conditions, which in turn correspond to diverse dynamics.[13]

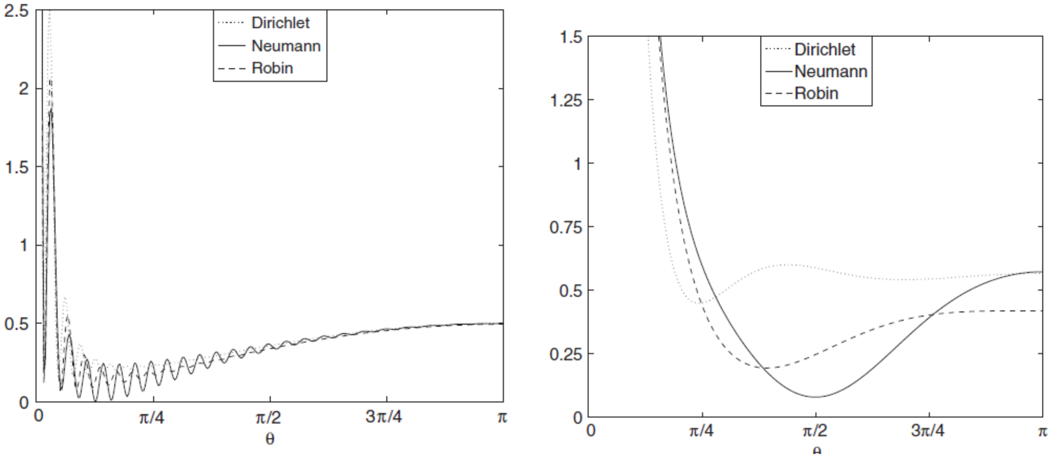

Fig. 3 Predicted differential cross section as a function of θ in the case of high energies (left) and in the case of intermediate energies (right), using Dirichlet, Neumann, and Robin boundary conditions. From de Oliveira & Pereira (2010, 25-27).

Three boundary conditions that are regularly discussed are known as Dirichlet, Neumann, and Robin (de Oliveira and Pereira 2010), and for the manifestation of the AB effect narrowly construed, any will do. However, the choice of boundary condition *matters* for, following Earman (2019), what we call the AB effect broadly construed, or *broad AB effect* for short. Particularly, imagine conducting scattering experiments wherein one scatters the electron beam off an infinitely long and absolutely impenetrable solenoid (as in I1-I3) and then inquiring into what happens when the solenoid is turned "off" versus "on." Again, classically we don't expect any special effect to manifest but according to quantum mechanics the deflection of the beam of electrons will be affected by the completely

---

[12] There is an essentially self-adjoint Hamiltonian operator representing the non-idealized case and corresponding to unique dynamics. See Appendix.

[13] There is an (infinity-fold) infinity of self-adjoint extensions of the Hamiltonian operator corresponding to the idealized case. See Appendix.

shielded magnetic flux. Crucially, the formal implementation of I1-I3 via a choice of boundary condition will gives rise to separate dynamical evolutions with *distinct empirically confirmable predictions* for such scattering experiments. See Fig. 3, which portrays graphs representing the "differential cross section" (which basically describes the collisions and deflection of the electrons off the solenoid) predicted by quantum mechanics for scattering experiments with different formal implementations of I1-I3 via Dirichlet, Neumann, and Robin boundary conditions.

Hence, the AB effect broadly construed actually concerns a family of AB effects predicted by quantum mechanics, corresponding to different boundary conditions. Moreover, the conditions that are relevant in a particular laboratory setting may be decided empirically: "in principle it is not obvious which boundary conditions are naturally realized in laboratories, and we have found that … *it is always possible* to find a range of the energy so that the corresponding scattering cross sections can be *distinguished*" (de Oliveira & Pereira 2010, 3; my emphasis). This means that in an analogous manner to the narrow AB effect, the abstract AB effect broadly construed has a concrete and empirically confirmable counterpart. There are thus four AB effects: abstract-narrow, concrete-narrow, abstract-broad, and concrete-broad.[14]

Historically, such distinctions were not introduced by Aharonov and Bohm (1959). Their paper, and the alleged experimental confirmation of the AB effect by Chambers (1960), were met with skepticism and resistance, leading to a thirty-year heated controversy in the physics literature. There were at least four main issues that were concurrently debated: (1) One concerns whether AB effect is a real, empirically verifiable effect that has been confirmed by laboratory experiments, and another has to do with (2) what exactly is the AB effect in the first place. (3) Another point of contention concerned the fundamental, ontological causal-mechanism that brings about the AB effect, e.g., is it the magnetic field/flux or a magnitude known as the electromagnetic potential (which is generally regarded as mathematical fluff without physical significance)? (4) Depending on one's stance regarding (1)-(3), another issue has to do with the foundational implications for quantum mechanics in light of the AB effect, e.g., does the AB effect signify a novel kind of indeterminism or non-locality (action-at-a-distance)?

For instance, in their paper titled "Nonexistence of the Aharonov–Bohm Effect," Bocchieri and Loinger (1978, 475) argued that the AB effect has a "purely mathematical origin," and that their analyses "leave no room for the effects of the kind of Aharonov's and Bohm's" (478). In reply, Klein (1979) and Bohm and Hiley (1979) responded in defense of Aharonov and Bohm (1959) noting flaws in Bocchieri and Loinger's work. Issues such as whether "there [is] an AB effect if the infinite solenoid is—as AB assume—absolutely impenetrable" were at the front and center of the debate (Bocchieri and Loinger 1981, 168). For example, Bocchieri and Loinger (1981, 168) argued that:

---

[14] Where the broad effects further concern families of AB effects.

> If the solenoid is turned on first, and a finite potential barrier is subsequently allowed to become infinite, an AB effect will be observed. [But the effect arises] from the penetration of the particle into the region in which [the magnetic field is non-zero] during the time in which the potential barrier is still finite. [On the other hand, if] one were to imagine that an impenetrable barrier had somehow erected first, and that the solenoid was turned on later, no AB effect would be observed.

The question about the reality of the effect was intermingled with its causal-mechanistic basis vis-à-vis the so-called electromagnetic vector potential—as the title of the original Aharonov and Bohm (1959) paper suggests: "Significance of electromagnetic potentials in the quantum theory." Hence, a brief interlude concerning said potential is warranted. In particular, in classical electromagnetism, electric and magnetic fields (and their respective fluxes) are considered to have genuine physical significance—these are things that exist in the world. It turns out though that one can represent a magnetic field with another mathematical entity, the vector potential, and that the relation is one to many. Namely, many different vector potentials can represent the same magnetic field. Classically, we take this to mean that not all degrees of the vector potential have physical significance—the many vector potentials corresponding to a particular field are just different ways of describing the same physical magnetic field. Roughly, foreshadowing discussion to come, choosing a particular potential corresponds to picking a "gauge," and any transformation of the potential that leaves the magnetic field invariant is a "gauge transformation," so that magnetic field is a "gauge-invariant" quantity.

Looking to Fig. 2, what Aharonov and Bohm (1959) noticed is that, while the completely shielded magnetic field (symbolized by $\mathbf{B}$) can only act on the electron beam via some non-local, (perhaps) spooky action-at-a-distance, there is a corresponding electromagnetic vector potential (symbolized by $\mathbf{A}$) that can mediate a local interaction. They also noted the corresponding Schrödinger equation is expressed with the vector potential. Thus, they suggested that in quantum mechanics the electromagnetic vector potential has physical significance:

> [I]n quantum theory, an electron (for example) can be influenced by the potentials even if all the field regions are excluded from it. … It would therefore seem natural at this point to propose that … the fundamental physical entities are the potentials, while the fields are derived from the them by differentiation. (Aharonov and Bohm 1959, p. 490)

Part of the ire then that Aharonov & Bohm (1959) drew has to do with this particular suggestion and, indeed, it would imply that quantum mechanics manifest an extreme type of indeterminism (even without the well-known "collapse of the wave function"). Consider, for instance, Magi and Valz-Gris (1995, 186; my emphasis) on the topic:

> Finally, the reader might wish that a definite answer be given to the vexed question about the existence or nonexistence of the Aharonov–Bohm effect. In order to answer it we have to split the question in two: (a) If, under the name of the Aharonov–Bohm effect a shifting of interference pattern is meant, then our answer

> is that, at any rate, the Schrödinger equation leads exactly to this prediction. (b) If, on the other hand, what is meant is a shifting *due* to any independent effect of the vector potential then our answer is no, there is no such thing coming from the Schrödinger equation.

The above citation also illustrated that the historical controversy around the AB effect did not only surround its experimental verification, causal-mechanistic basis, and foundational implications; another main issue concerned what the AB effect actually is *in the first place*. Do we mean an effect that is due to a completely shielded magnetic field/flux? Or due to the electromagnetic vector potential? Is the effect purely quantum mechanical in origin? Or does it have classical counterpart? Is it a topological or dynamical effect? And so on.

For instance, say that one characterizes the AB effect as an effect that is *due to the electromagnetic vector potential*, (as the title of Aharonov and Bohm (1959) misleadingly suggested). Earman (2019, 2011) notes how "skeptics [in turn] demanded that for actual experiments to count as confirmation of the AB effect there should be no plausible way to attribute the actually observed effects to" an interaction of the electron beam and magnetic field in a region accessible to both. Strocchi and Wightman (1974) then argued that, since in reality there are no *completely* shielded magnetic fields (as I1-I3 suggest), there will always be some minimal interaction between the electron beam and the field. Thus, they claimed, no actual experiment can ever confirm the AB effect, if what is meant by the effect is one that is *due* to the electromagnetic vector potential.

Moreover, going back to the issue of the reality of the effect and its experimental verification more generally speaking, one may wonder: If the AB effect was predicted in 1959 by the work of Aharonov and Bohm (1959), or perhaps as early as 1949 by Ehrenberg and Siday (1949), and subsequent experimental test began as early at 1960 with Chambers (1960), why is it that the physics community didn't come to accept the effect as a bona fide physical effect until the experiments of Tonomura *et al*. (1982, 1986)? Why was there such a vibrant controversy surrounding the effect for about thirty years, with deniers existing up until this day, e.g., Boyer (2000a, 2000b, 2006, 2008), Bunge (2015), and Wang (2015)? Earman (2019; 1992, 2010) argues that (at least part of) the answer has to do with the fact that the abstract AB effect cannot ever be manifested in the actual world so this leads to the following puzzle:

> [H]ow can experiments on actual systems serve to confirm the predictions of the theory for the behavior of the fictional system? What conditions must an actual world apparatus satisfy in order that it can produce confirmation of the AB predictions for the fictional apparatus?

It is worthwhile to pause to consider how odd the situation is. The issue is not about whether experiments can confirm an interference shift due to a magnetic field—they can, and even classically we expect such a shift. Rather, the issue concerns which *actual* experiments can confirm predictions of quantum theory about a *possible idealized* scenario (as in I1-I3), viz., the abstract AB effect. Moreover, say, it is agreed that experiments have come close enough to the ideal to confirm the presence of the abstract AB effect. We can then reinterpret the interference shift that we observe in the laboratory as the physical counterpart of the idealized effect, viz., as the concrete AB effect. In contrast, if the abstract AB effect didn't manifest in the idealized scenario, then there is no reason to think that the observed interference shifts concern some novel (worthy of its own name) "effect." Consider Tonomura *et al*. (1986, 794) on the matter:

> The most controversial point in the dispute over experimental evidence for the AB effect has been whether or not the [interference] shift would be observed when both electron intensity and magnetic field were extremely small in the region of overlap. Since experimental realization of absolutely zero field is impossible, the continuity of physical phenomena in the transition from negligibly small field to zero field should be accepted instead of perpetual demands for the ideal; if a discontinuity there is asserted, only a futile agnosticism results.

To end, the historical controversy surrounding the AB effect shows that in thinking about the idealizations involved in the AB effect, viz., the abstract AB effects, physicists were concerned with issues such as (1) the reality and confirmation of the effect, (2) its nature, (3) its ontological, causal-mechanistic basis, and (4) the corresponding foundational issues. Very generally then, I will take "the AB effect" to refer to the quantum mechanically predicted existence of some observable effects in electron beam that reflects the strength of the magnetic flux inside the solenoid—while taking a quietist stance on foundational and ontological issues. But the reader will recall that I have identified four AB effects (abstract-narrow, concrete-narrow, abstract-broad, and concrete-broad), and so by "AB effect" I am referring to all four AB effects.[15] In the abstract cases, "observable" effects are predicted but cannot be directly confirmed. They may be indirectly confirmed through their concrete counterparts, but (as noted in the Earman quote above) what is needed for such experimental confirmation is controversial. In what follows, I aim to clarify in what sense the idealizations involved in the AB effect—as manifested in the abstract AB effects—afford understanding, with an eye primarily to the notions of understanding-with and understanding-what.

4. Understanding-with: Objectual understanding of scientific theory.

Reflecting on the AB effect case study and controversy suggests that three related factors are indicative of understanding-with, viz., objectual understanding of a theory like quantum mechanics, including (but not limited to) (i) knowing a theory's the modal structure, (ii) knowing a theory's foundations (e.g., is the theory deterministic/indeterministic, local/non-local?), and (iii) knowing the (relevant) intertheoretical relations (of said theory). In what

---

[15] Similarly, by "abstract AB effects" I'm referring to both narrow and broad abstract AB effects, by "concrete AB effects" I'm referring to both narrow and broad concrete AB effects and so on.

follows I consider each in turn (in reverse order), and then in the following section I situate my discussion in relation to philosophical accounts of understanding.

Beginning with (iii), the language used by proponents of the AB effect controversy suggest an important contrast class, namely, how quantum mechanics (or a particular interpretation) relates to predecessor and competing theories. For instance, Aharonov and Bohm (1959, 490) note that "In classical mechanics … the [electromagnetic vector] potentials have been regarded as purely mathematical auxiliaries, while only the [electromagnetic] field quantities were thought to have a direct physical meaning." They then continue to contrast the classical situation with what they have argued happens in quantum mechanics, viz., the abstract AB effect is predicted. Notice how the abstract AB effect, along with the idealizations I1-I3 involved, are not appealed in order to simplify a formal representation of some actual experimental setup. Nor are they used to flag "difference-makers" (relevant factors) and/or irrelevant factors for the empirical manifestation of the AB effect (viz., the concrete AB effect). Such use of idealizations diverges from the ways philosopher usually describe them. For instance, Strevens' (2017) holds that idealizations "help us understand a phenomenon by giving us a better grasp of a correct explanation of the phenomenon" (Strevens 2017, 42). He takes idealized models to be favored over non-idealized models because (a) such models are simple, practical, and straightforward, (b) logically or mathematically simple, and (c) idealizations offer salient suggestions and tractable frameworks for exploring various derivations (or attempted derivations) "of the phenomenon to be explained, derivations that differ with respect to which causal factors are present and which are not" (46-47).

However, (a)-(c) are *not* importantly applicable to the AB effect case study. Instead, the abstract AB effect makes salient the radically different structure of quantum and classical mechanics, via their predictions about the behavior of an electron beam in the vicinity of a *completely* shielded magnetic field. Alternatively, if we de-idealize so that the magnetic field can interact directly with the electron beam then *both* quantum and classical physics predict shifts in interference pattern. One may object that even in the context of the de-idealized concrete AB effect, the predicted quantum and classically shifts will differ. In reply, note that some differences are not substantive because they are differences of degree: both quantum and classical treatments predict shifts in interference pattern but the shifts differ slightly. Instead, the difference between the quantum and classical treatment of the idealized scenario varies dramatically and qualitatively: quantum mechanics predicts the abstract AB effect and classical physics predicts that there will be no effect whatsoever.[16] To stress this point, imagine that both quantum and classical physics predicted slightly

---

[16] Admittedly, this claim is ostensibly in tension with talk of the "classical analogue" of AB effect and the idea that we can, classically, expect some type of shit in real experiments. To clarify then, the "classical analogue" of the effect concerns the general theory relativity so this effect is not directly related to our discussion. As for the shift in interference patterns, in the idealized scenario that we are discussing, classically, no shift is predicted. In real experiments we expect there to be some shift due to the classical electromagnetic interaction between the beam of electrons and the shielded magnetic field. Still, even such behavior presupposes those electrons has a wave-nature (which is not, strictly speaking, classical).

different phase shifts in the presence of a magnetic field such that when the magnetic fields is *completely* shielded, *both* predict *no* shift. Had things turned out this way, there would be no effect named the "AB effect."

Consider another example focusing on (iii). As is well known, quantum mechanics has various competing interpretations, and these can be viewed as competing theories (since they differ in their ontological implications and some are not even empirically equivalent). One way to assess the viability of a particular interpretation would be to note its implications for the AB effect. For example, if orthodox (Copenhagen) quantum mechanics predicted the AB effect in such a way that implied unreasonable indeterminism and/or non-locality, but, say, Bohemian quantum mechanics could supply a deterministic and local account of the AB effect, this may bode well for the latter interpretation. Such an investigation was undertaken by Healey (1997), who uses abstract AB effect to examine competing quantum interpretations. He argued that on various interpretations of quantum mechanics, the AB effect manifests non-locality either by a violation of the "principle of local action, or of a principle of separability, or of both" (39). The exact details need not concern us here,[17] but claims to the effect that "…a strong version of the Copenhagen interpretation… does not conform to *Local Action*…" (32-33), "…the weak version of the Copenhagen interpretation… [is] in conformity to *Local Action* [but not to separability]" (36), and that (depending on the details) the Bohmian interpretation violates both local action and separability (36-39), show how the abstract AB effect can sheds light on competing interpretations of quantum mechanics.

Similarly, Earman (2019, 2016) claims that a more felicitous setting for resolving issues about locality versus non-locality may be by comparing the abstract AB effect in quantum mechanics, which suggest that gauge invariant observables (like the magnetic field) are non-local, with its manifestation (or lack thereof) in quantum field theory. Likewise, Madulin (1998, 367; my emphasis) notes:

> In the end, the Aharonov-Bohm effect points us to the very important problem of *understanding* the physical significance of quantities which admit of local gauge transformations, indeed of *understanding the meaning* of gauges and gauge transformations *at all*. Since local gauge-invariance has become a cornerstone to contemporary quantum theory, this investigation is central to the ontological interpretation of the quantum theory…

Setting details aside, what is crucial here for my purposes is that the abstract AB effect plays a role in helping us understanding concepts that are "a cornerstone to contemporary quantum theory."

---

[17] But, for the interested reader, the difference between "local action" and "separability" is as follows (Healey 1997, 23-24): "*Local Action*: If A and B are spatially distant things, then an external influence on A has no immediate effect on B" and "*Separability*: Any physical process occurring in spacetime region R is supervenient upon an assignment of qualitative intrinsic physical properties at spacetime points in R."

Moving on to (ii), foundational issues already arise in the discussion of intertheoretic relations through both the controversy surrounding the ontological implications of the effect vis-à-vis its causal-mechanistic basis, and the discussion of the meaning of gauges and gauge transformations in quantum theory. A further example is exhibited by the Healey-Maudlin debate (Healey 1997, 1999; Maudlin 1998) regarding the foundational implications of the abstract AB effect for quantum theory. In particular, Healey's (1997) main point is that there is an interesting analogy between the non-locality that arises in the abstract AB effect and quantum entanglement:

> The kind of nonlocality manifested by the Aharonov-Bohm effect is much more closely analogous to the kind of nonlocality manifested by violations of Bell inequalities than has been previously acknowledged. Neither effect can be given a completely local explanation. But in both cases one may analyze the residual nonlocality as involving the violation either of a principle of local action, or of a principle of separability, or of both; and in both cases, exactly how one analyzes the nonlocality depends on how one interprets quantum mechanics. (Healey 1997, 39)

Maudlin (1998), however, rejects the foundational import that Healey extracts from said analogy:

> This conclusion may come as something of a shock to those who have studied the Bell inequalities, for one would suppose, on fairly obvious grounds, that the Aharonov-Bohm effect can be given a completely local explanation, which violates neither local action nor separability … The puzzles about the Aharonov-Bohm effect lie in the interpretation of the quantum formalism, not in the impossibility of providing a local account (in [John] Bell's sense). Maudling (1998, 362-363)

For our purposes, what is important to note is the role that the abstract AB effect is playing this debate: it is shedding light on foundational issues of quantum mechanics having to do with locality, the nature of locality, and the relation to other types of non-locality in quantum mechanics via violations of Bell's inequalities. In fact, we can add that an important disanalogy between the abstract AB effect and violations of Bell's inequalities is that the source of non-locality in the former concern observables like the magnetic field, while the latter concern the quantum state itself. The point though is that such foundational issues arise, and conclusions for quantum mechanics can be investigated, specifically in the realm of the abstract AB effect.

Finally, it is by looking to (i) the modal structure of quantum mechanics, and specifically to what it says about possible worlds that are most naturally construed as idealizations (vis-à-vis the abstract AB effect) that we facilitate (ii) and (iii). This point ought to be highlighted in order to evade two possible charges of triviality. The first notes that non-actual but possible worlds can't be idealizations since this would imply that any non-actual values of variables in even the most mundane of cases would be idealizations. The second holds that, if a theory involves idealizations, then it is trivially true that idealizations are necessary in order to gain understand of said theory; and the same is true of gaining modal understanding of possibilities. I believe though that these worries are unfounded for the following reasons.

Consider Chambers' (1960) original experiment, which was conducted in order to confirm the concrete-narrow AB effect, and had a set up similar to Fig. 2, only with finite sized solenoids, etc. This is an actual experiment that occurred in our world—it is not a world naturally construed as an idealization even though when we use quantum mechanics to represent and make predictions about Chambers' experimental setup we'll idealize or abstract away irrelevant details, use approximations, etc. A different type of experiment that is closer to the ideal setup of the abstract AB effect occurs is a possible world wherein we use a toroidal solenoid covered with a superconducting film (since such a setup is better at keeping the magnetic field shielded from the electron beam and vice versa). This is also not a world naturally construed as an idealization since it is a possible world that could have been actualized and, in fact, was actualized in the Tonomura *et al*. (1986) experiment.

In contrast, the abstract AB effect, being a prediction of the Schrödinger equation, concerns a possible world (according to quantum mechanics) that cannot be actualized due to contingent, pragmatic constrains: there isn't enough time, space, energy, or money, to build an infinitely long and absolutely impenetrable solenoid. Moreover, *this* possible world is naturally construed as an *idealization* due to the presence of I1-I3, resulting in a *completely* shielded magnetic field. In other words, concerning the first charge of trivially, it is true that scientist appeal to possible but non-actual worlds all the time and that this doesn't mean that idealizations are essential for understanding-with. However, not all possible but non-actual worlds are idealizations, and not all such idealizations are used by scientist to shed light on (ii)-(iii) as happens in the AB effect.

As for the second charge of triviality, it is true that theories, generally, involve idealizations and that modal understanding involves appealing to possible but non-actual worlds. That's trivial. What isn't trivial is that a possible but non-actual world that is naturally construed as an idealization (e.g., because it involves infinite solenoids and potential barriers) plays a role in bringing about understanding-with via (ii)-(ii) and understanding-what (as argued below) in the context of an actual document historical controversy.

5. An Outline of An Account of Understanding-with.

In this section, I will situate the above in the context of the understanding literature, as well as outline an account of understanding-with that is suggested by reflection on the AB effect case study. Starting with Khalifa (2017), recall that he takes understanding-why p to be constituted by the relevant explanatory knowledge related to *empirical* phenomenon p via his EKS account. Moreover, Khalifa's notion of understanding admits of degrees so that one's understanding of p improves as one grasps a greater number of correct explanations of p and as those explanations bear a closer resemblance to scientific knowledge. A natural and minor extension of Khalifa's account to an account of understanding-with would hold, along the lines of Strevens (2013), that understanding a theory is constituted by the relevant explanatory knowledge related to all p that are within the purview of said theory; and that understanding-with similarly admits of degrees. The problem is that such an account doesn't accommodate (i)-(iii) since, while the concrete AB effect is an empirical phenomenon p that is governed by quantum mechanics, the abstract AB effect is, well, abstract. Moreover, even if we extend the suggested account to include all phenomena (concrete-empirical and abstract) governed by quantum mechanics, thereby catering to (i), it isn't clear that "explaining" such phenomena—which more or less follows directly from Schrödinger's equation—will give us knowledge about (ii) and (iii). Whether such knowledge is forthcoming depends on one's account of explanation.

Since the notion of explanation is central to various accounts of scientific understanding (e.g., Khalifa (2017), De Regt (2017), Strevens (2013)), it is worthwhile to briefly discuss what explains the AB effect. Setting aside the controversy regarding experimental confirmation and the justification of details concerning boundary conditions, recall that Aharonov and Bohm (1959) derive the abstract AB effect from the Schrödinger equation. Hence, if one is inclined toward a nomological account of explanation (e.g., Hempel and Oppenheim (1948)), and one is willing to grant the Schrödinger equation the status of a "law of nature," then the AB effect is explained in the sense that it can be deduced from the Schrödinger equation once the details of the setup are taken into account. Instead, if one is more partial to difference-making (Strevens 2008, 2013; Khalifa 2017) or counterfactual (Woodward 2003) accounts of explanation, one can also explain the AB effect since a linear partial differential equation like the Schrödinger equation affords counterfactual inferences (including inferences corresponding to relevant difference-making factors). For instance, if we are willing to consider the abstract-narrow AB effect setup as a kind of *idealized model*, we can use it to answer Woodward's what-if-thing-had-been-different questions (*w-question)* and we can (partially) *justify* answers via rigorous mathematical results (e.g., de Oliveira and Pereira (2008, 2010, 2011), Ballesteros and Weder (2009, 2011)) that lend support to the physical manifestation of the concrete AB effects. This means the abstract-narrow AB effect setup satisfies Bokulich's (2008) model-explanation account. And if one is satisfied by such explanations, we have some explanatory understanding of why the concrete AB effects occurs that caters to (i), but not to (ii) and (iii).

Instead, if one insists on a causal-mechanistic explanation of the AB effect, then there will certainly be clear implications for (ii) and (iii). However, it isn't clear anymore that we have an explanation of the AB effect. The reason, in short, is that the causal-mechanistical basis of the AB effect is an issue of current controversy, analogous to debates about the one true interpretation of quantum mechanics. Interpretative options include a non-local (gauge-invariant) magnetic field, a local physically significant (gauge-dependent) potential (Aharonov & Bohm 1959), additional field like entities that restore locality (Mattingly 2006), physically significant (gauge-invariant) holonomies (roughly, properties attributed to sets of possible loops around the solenoid) (Belot 1998, Healey 2007), representing the entire setup with the quantum wave function (Vaidman 2012), and other (allegedly) local and gauge-invariant interpretations (e.g., Wallace 2014). Thus, on a causal-mechanistic view (ii) and (iii) are naturally attended to but there is no non-controversial straightforward explanation of the AB effect. In contrast, on a non-causal-mechanistic view we have a straightforward explanation but (ii) and (iii) are missing from our modified Khalifa-understanding-with account.

In principle, similar comments to those made regarding Khalifa's (2017) account can be made regarding De Regt's (2009, 2017) CUP account when modified (in the ways suggested above) to accommodate the notion of understanding-with. However, De Regt (2009, 30-31) considers a type of pragmatic notion of understanding-with: "Pragmatic understanding … is defined, somewhat loosely, as the ability to use the relevant theory and as being based on skills and judgment" (also see De Regt 2017, 90-92). He develops this idea with the concept of "intelligibility," which is also central to his CUP account. Namely, intelligibility is "the value that scientists attribute to the cluster of qualities [e.g., simplicity, continuity, scope, familiarity, appeal to causes, unification, mechanism, visualizability] of a theory (in one or more of its representations) that facilitate the use of the theory (De Regt 2017, 40). A sufficient condition for intelligibility, e.g., concerns scientists' ability to "recognize the qualitatively characteristic consequences of [a theory] without performing exact calculations" (102). One could then accommodate (i)-(iii) by integrating such factors into the notion of intelligibility, noting that the AB effect case study shows how scientists attribute value to knowing (i)-(iii). More recently, De Regt and Gijsbers (2017, 55-56) have argued that "…understanding can only be gained from representational devices that are" effective in the sense that they can be reliable used by a scientist to make correct prediction, guide practical applications, and develop better science. At least the last of these criteria seems to accommodate (i)-(iii), so it is plausible to generate an account of understanding-with based on such work.

Another approach that is worthwhile to consider since it specifically targets the notion of "understanding a theory" is Newman's (2017, 582) IMU. Generally, if knowing (i)-(iii) concerns "qualitative problems relevant to" a theory of interest, then IMU does justice to case studies like the AB effect. Looking to Newman's (2017, 580-581) discussion of the type of problems that are of interest—some examples include dropping a ball, submerging a balloon tied into water, the electrical force between an electron and a proton, the sun collapsing into a blackhole near the Earth, and a dive bomber releasing a bomb when diving, climbing, or flying horizontally—one gets the feeling that there is some dissonance between Newman's interests and mine. This ought to already be clear at the start of Newman's (2017, 578-59) paper, where he states: "What does it take to display that one understands, say, introductory quantum mechanics? Surely passing exams is a big part of the answer."

Perhaps understanding quantum mechanics is similar to understanding introductory quantum mechanics. The idea would be that, for instance, the abstract AB effect affords a "simple pattern of reasoning," wherein we apply the principles of quantum mechanics using a "mental model," and the solutions avails itself (581). The fact, however, is that deriving the AB effect, in any of its manifestation, let alone understanding quantum theory in light of the AB effect, is immensely more complicated. As Earman (2019, 1994) notes: Part of what is "missing from the philosophical literature is any sense of just how subtle, complex, and rich an analysis is required for the fictional systems that display the AB effect." It is hard to see how the motivations leading Newman to his IMU—for example, taking his lead from Paul Hewitt's 'Conceptual Physics' webpage[18]—can do justice to the type of intricate historical controversy surrounding the AB effect. In any case, if by "qualitative problems" Newman does have in mind (i)-(iii), then my analysis here can be interpreted to fill in some details on how understanding-with manifests in the AB effect in context of his theory.

In sum, perhaps both Khalifa's (2017) and De Regt's (2017) accounts could be modified in ways to accommodate (i)-(iii) and a theory of understanding-with, and maybe Newman's (2017) account already fits the bill. Still, much of the understanding that the abstract AB effect affords of quantum mechanics is powered by looking to the modal structure of quantum mechanics. This suggest that modality ought to take center stage in a discussion of understanding-with. Such an approach is provided by Levy (2019, 282), who states that for "Subject S to understanding target T amounts to S's being able to use a representation R of T in order to draw sufficient accurate and general inference about T's (actual and) counterfactual behavior." Also, Le Bihan (2017, 112) recent account of "modal understanding" holds that: "One has some modal understanding of some phenomena if and only if one knows how to navigate some of the possibility space associated with the phenomena." A natural extension of such approaches holds that one has *understanding-with* of a theory if one knows how to navigate its possibility space, noting the particular importance of possible worlds that illuminate (ii) and (iii). To that effect, reflecting on the AB effect case study and the historical controversy, I submit that it suggests an outline of an account of understanding-with along the following lines:

> Subject S has understanding of a theory T to the extent that S is able to (i) draw accurate and general inferences about phenomena governed by T and navigate the modal space associated with T (i.e., associated with the possible worlds according to T), (ii) answer various foundational question about T (such as whether T is deterministic/indeterministic, local/non-local, supports Humean supereminence or holism/emergence, supports or rejects commonsensical views of space, time, objects and properties, causation, etc.), (iii) elucidate the relations of T to predecessor and competing theories and interpretations.

On the face of it, such an account seems to set the bar too high since, e.g., even experts on quantum mechanics may lack knowledge about its foundations. However, understanding-with is not an all or nothing affair. Rather, understanding-with is a matter of degree, and the more you know about the modal and foundational structure of, and intertheoretic relations corresponding to, a theory, the more you understanding the theory. Consequently, only a very naïve conception of understanding-with can dispense with idealizations given the essential role they play in affording modal, intertheoretic, and foundational information.[19]

---

[18] http://www.arborsci.com/60-questions-physics-students-should-know

[19] Compare with Karl Popper (1972, 299): "[A] full understanding of a theory would mean understanding all its logical consequences. … But this means that nobody, neither its creator nor

To end, let me remind the reader of the scope of my claims here, as well as what further work is suggested. I am claiming that any theory of understanding-with that will do justice to the AB effect case study will have to incorporate (i)-(iii) as factors that are (at least) indicative of understanding a scientific theory. This suggests an outline of a theory of understanding-with that puts modality at front and center, as noted above. Having said that, I am not claiming that (i)-(iii) are constitutive of understanding-with; that they are necessary and/or sufficient conditions for objectual understanding of a theory. Such a claim would necessitate a greater defense than I can muster here. Thus, the development of my suggested outline, a comparison with competing accounts, and an identification of merits and faults must await another occasion.

One may wonder though whether there are other case studies that support thinking about understanding-with in the manner that I have suggested. If not, it would mean that the AB effect is an anomaly; it is specific and it's hard to see how to generalize my claims to other examples. Unfortunately, it is beyond the scope of this paper to attend to other case studies in any detail, as such analysis is complex and nuanced (as I hope is clear from my discussion of the AB effect). What I can do here is gesture at how other examples my fit the bill. Starting with a toy example, Norton's Dome (Norton 2008) concerns a non-actual but possible world according to classical mechanics, which cannot be actualized since the curvature diverges at the Dome's apex (Malament 2008), and is thus naturally construed as an idealization. Moreover, it sheds light on a foundational issue in classical mechanics, viz., indeterminism. Still, the Dome is neither an actual example from the history of science wherein a controversy revolved around (i)-(iii), nor does it correspond to a concrete, physical counterpart.[20]

There are three other examples that are worthwhile to mention. First, (truly) thermodynamic reversible processes correspond to idealizations that play central roles in thermodynamics and statistical mechanics (Ladyman (2008), Valente (2019)). One can also appeal to an idealized scenario, vis-à-vis "Maxwell's Demon," to shed light on foundational and intertheoretic issues, namely, how statistical mechanics differs from thermodynamics in that the former's version of the second law is statistical so that "we come to understand the meaning or semantic content (including modal strength) of Maxwell's statistical second law" (Stuart 2018). Second, Shech and Gelfert (2019) discuss how idealizations (in the form of idealized limit cases) facilitates modeling the Mott phase transition with the Hubbard model. In particular, exploring the modal structure of the Hubbard model via idealized limiting procedures affords an understanding of the nature of the Mott metal-insulator transition allowing us to answer foundational questions such as whether such transitions are thermodynamic or quantum in origin, if quantum then whether they are discontinuous or continuous, whether there is a line of first-order transitions at finite temperatures, etc.

---

anybody who has tried to grasp it, can have a full understanding of all the possibilities inherent in a theory…"

[20] Not for lack of trying: http://www.pitt.edu/~jdnorton/Goodies/dome_cam/dome_cam.html

Last, idealizations in the form of the two-dimensional limit allow for the exploration of full spectrum of the quantum exchange phase and corresponding quantum statistics (Shech (2015, 2019), Shech and Gelfert (2019), Shech and McGivern (2021)). Specifically, in two-dimensions quantum theory has the representational capacity to characterize anyons, which are particles that obey fractional quantum statistics and are contrasted with bosons and fermions (Wilczek (1982), (1990)). Fractional statistics in turn play are role in explaining the fractional quantum Hall effect (Bain (2016), Shech and McGivern (2021)). In all such cases, idealizations facilitate exploration of modal structure, which in turns affords understanding along the lines of (i)-(iii) discussed above. But arguing for this claim is beyond the scope of the current paper. Still, idealizations also play a central role in part of the characterization of phenomena like thermodynamic reversible process and phase transitions, and entities such as anyons. This is related to my discussion of the notion of understanding-what a phenomenon is in the first place, which I turn to next.

6. Understanding-what: Objectual understanding of phenomena.

Part of the historical controversy surrounding the AB effect, its foundational implications, and its experimental verification, had to do with the fact that it was not always clear what the AB effect what supposed to be the first place. Historically, Aharonov and Bohm (1959) characterized the AB effect along the lines of what I have been calling the abstract AB effect and discussions of the effect in much of the theoretical physics and philosophy of physics literature, as well as textbook accounts, use this conventional definition (usually focusing on the narrow-abstract AB effect). Aharonov and Bohm also suggested that there is a physical, empirically confirmable manifestation of the effect, viz., the concrete AB effect. Hence, taking lead from the convention, one could characterize the concrete AB effect as the physical counterpart of the abstract AB effect. That is to say, what we expect to observe in the laboratory (given various rigorous results, e.g., de Oliveir and Pereira (2008, 2010, 2011) Ballesteros and Weder (2009, 2011)) when we de-idealize from I1-I3. In this sense, reference to the abstract-narrow AB effect is essential to understanding-what the concrete-narrow AB effect is. But one may wonder, is it necessary? What is the problem of characterizing the AB effect as a concrete, physical phenomenon, eschewing any reference to the idealized, abstract AB effect? Let me first discuss the narrow AB effect, and then broad AB effect.

One problem with characterizing the concrete-narrow AB effect as shift in interference pattern due to a *partly* shielded magnetic field is that, classically, we expect some kind of shift due to the (classical) interaction between the electron beam and magnetic field. There is nothing novel or surprising about this type of effect. As I noted above, quantitatively, the predicted shift between the classical and quantum case will differ slightly. However, characterized in this manner, the difference isn't a categorical and qualitative one as the historical controversy suggests, it is just one of degree. This motivates the following qualification: a shift in interference patterns due to a partly shielded magnetic field that cannot be explained non-quantum mechanically (e.g., classically, relativistically). In other words, the effect must be a genuine, independent quantum effect. I find this suggestion rather odd because it characterizes the concrete, narrow AB effect negatively—"an effect that can't be explained classically." Furthermore, arguably, any shift in interference pattern due to an almost shielded magnetic field can be explained (away) classically. This is part of what led to the thirty-year controversy surrounding the AB effect (as noted in Section 3), and it sustains the controversy until this day among some.

For instance, in a series of papers (e.g., Boyer (2000a), (2000b), (2006), (2008)), Timothy Boyer has argued that there is an additional induced electromagnetic force between the electron beam and the solenoid involving "subtle relativistic effects, " which displaces the solenoid center in a manner that accounts for the shifted fringe pattern seen in experiments allegedly confirming the concrete-narrow AB effect. Boyer notes that "at present the only experimental evidence we have is the Aharonov-Bohm [interference] shift itself, and this [interference] shift is indistinguishable from those due to classical electromagnetic forces" (2000a, p. 898). In addition, some (e.g., Lyre (2009)) hold that the AB effect is not quantum mechanical in origin because of the analogue classical gravitational AB effect. Such views substantiate the following conjecture: without a reference to the abstract-narrow AB effect, one can, in principle, always explain away the concrete-narrow AB effect, as a non-quantum mechanical effect. Such a line of reasoning suggests that the abstract-narrow AB effect is not only essential for understanding-what the concrete-narrow AB effect is, it may be necessary.

In order to bolster this last claim, consider the context of the concrete-broad AB effect where there is something like an in-principle argument that can be given for why reference to the abstract-broad AB effect is necessary for understanding-what. Recall, the emergence of the abstract-broad AB effect comes about from the fact that there are different formal implementations of I1-I3, corresponding to different boundary conditions obtained at the solenoid border (viz., Dirichlet, Neumann, Robin), which in turn correspond to different predicted dynamics in (say) scattering experiments. We can then characterize the concrete-broad AB effect as the physical counterpart of the abstract-broad AB effect, what we expect to observe in the laboratory when we de-idealize a bit from I1-I3 given the work of, e.g., de Oliveira and Pereira (2010) (see Fig. 3). Having said that, from a theoretical perspective, once we de-idealized, there is no longer an issue of formally implementing I1-I3.[21] In other words, understanding-what the concrete-broad AB effect is in the first place, even more so that the concrete-narrow effect, seems to necessitate an appeal to the abstract-broad AB effect.

To end, it would do well to position the above discussion of understanding-what with notions of objectual understanding of phenomena in the literature such as Kelp (2015)

---

[21] See Appendix. In detail, if we transition to the concrete setting where the beam of electrons can access the entire configuration space of $\mathbb{R}^3$, then the natural (initial) domain for Aharonov-Bohm Hamiltonian $H^{\boldsymbol{A}\infty}$ is $D(H^{\boldsymbol{A}\infty}) = C_0^\infty(\mathbb{R}^3)$ which is dense in $\mathcal{H} = L^2(\mathbb{R}^3)$. With this choice of domain, $H^{\boldsymbol{A}\infty}$ is essentially self-adjoint and thus there is no need for self-adjoint extensions corresponding to different boundary conditions.

and Dellsén (2020). Starting with Kelp (2015, 3813), his account of "Outright Understanding" holds that:

> [Subject] "S understands [phenomenon] P" is true in context c if and only if S approximates fully comprehensive and maximally well-connected knowledge of P closely enough to be such that S would (be sufficiently likely to) successfully perform any task concerning P determined by c, if, in addition, S were to have the skills needed to do so and to exercise them in suitably favourable conditions.

Accordingly, "S understands P" is true if S has "fully comprehensive and maximally well-connected knowledge of P," which, in turn, is equivalent to having "fully comprehensive knowledge of the *full account* of the phenomenon." "Full account" is the union of "of a phenomenon's description and story," where "description" is "the set of true propositions that describe" P and "story" is "a set of true propositions describing [P's] place in a broader nexus of phenomena" (Kelp 2015, 3809). But here's the rub: Kelp's (2015) notion of a phenomenon P, along with the broader nexus of phenomena appearing in P's story, concern *actual* phenomena: "Crucially, phenomena must be actual … [whereas objects, relations, and states of affairs that] do not exist at the actual world… do not count as phenomena" (3808).

This means that Kelps' (2015) account of objectual understanding of phenomenon cannot do justice to understanding-what the concrete AB effect is in the first place, since appealing to the non-actual but possible abstract AB effect is essential or even necessary. However, a minor adjustment to Kelp's (2015) account can remedy the problem: allow for both actual and possible phenomena to count as phenomena. It isn't clear to me what possible negative implications this would have for Kelp's account, if any; nor am I clear on the emphasis placed by Kelp regarding phenomena being "actual."

Next, Dellsén (2020, 1268) suggests a "dependency modelling account" (DMA) of objectual understanding of phenomena:

> DMA: [Subject] S understands a phenomenon, P, if and only if S grasps a sufficiently accurate and comprehensive dependency model of P (or its contextually relevant parts); S's degree of understanding of P is proportional to the accuracy and comprehensiveness of that dependency model of P (or its contextually relevant parts).

Can the DMA do justice the notion of understanding-what as it arises in the AB effect case study? Prima facia, if we take "P" to stand in for the "concrete AB effect," and if the concept of a "sufficiently accurate and comprehensive dependency model" is general enough to allow S to appeal to the abstract AB effect, then it seems we have a positive answer. Having said that, there are reasons to doubt that DMA naturally includes reference to non-actual but possible worlds, which are most naturally construed as idealization, as in the abstract AB effect.

For one, Dellsén's (2020, 1265) says that "a model is simply an information structure of some kind that is interpreted so as to represent its target" and that "[by] their nature, all models are incomplete or inaccurate representations of their targets." This suggests that Dellsén's (2020) notion of phenomena is similar to that of Kelp (2015), and so does not do justice to ideas manifested by AB effect case study. For example, Earman (2019, 1992) notes that "the center of interest is on effects that should, according to [some] theory, be exhibited by [a] fictional system." The "target system [of interest] in the AB effect is a fictional system," viz., the abstract AB effect, and the model of such an effect is "an accurate and precise description…" (1993). The point is that in Dellsén's (2020) (and Kelp (2015)) there seems to be no awareness of the fact that in order to understanding some phenomenon p (like the concrete AB effect) one may need to appeal to some other idealized phenomenon p' (like the abstract AB effect) and such "need" doesn't stem from issues of tractability or convenience.

Second, Dellsén's (2020, 1264) notes that "it is one thing to understand a phenomenon itself and quite another to understand the theories that may or may not be true of those phenomena," after which he refers to heat and the phlogiston theory of heat as an example illustrating this point. Fair enough, but some phenomena (like PT and the AB effect) do not have a reasonable theory-independent characterization. Again, this suggests that DMA is constructed to deal with phenomenon such that to "understand a phenomenon itself" can be clearly distinguished from understanding-with.

Last, Dellsén's (2020, 1267) claims that DMA "helps to explain why idealizations and abstractions can both increase understanding." He identifies the tension between sacrificing accuracy ("a kind of idealization") and comprehensiveness ("abstraction") as key: "Since understanding is a function of how well a model does on both accuracy and comprehensiveness at once, it is sometimes possible to increase one's understanding by sacrificing either accuracy or comprehensiveness…" (7). Generally, I have no qualms with Dellsén's claims in cases where idealizations are appealed to in the model of some concrete phenomenon for the purposes of making an intractable problem tractable, for flagging relevancies/irrelevancies, or making salient some aspect of interest such as visualizability. However, this isn't what is going on when one appeals to the abstract AB effect in order to understand-what the concrete AB effect is in the first place—it isn't an issue of sacrificing either accuracy and comprehensiveness.

In sum, although both Kelp (2015) and Dellsén (2020) accounts of objectual understanding of phenomena are (or with minor modification can be made to be) consistent with the AB effect case study, the details surrounding their account suggest a lack of awareness of the kind of objectual understanding of phenomena—understanding-what—that I have discussed here.

7. Recap & Understanding-why the AB effect Occurs.

In this section, I will summarize the claims that I have made in the context of the AB effect historical case study and quickly draw out implications for understanding-why the AB effect occurs. Relatedly, I will discuss worries to the effect that idealizations are neither necessary nor sufficient for understanding.

To begin, I claimed:

> C0: Idealizations are trivially necessary for understanding-what the abstract AB effect is in the first place.

This is so because the abstract AB effect is defined in reference to idealizations (viz., I1-I3).

> C1: Granting (i)-(iii), the abstract-narrow AB effect is essential for understanding-with quantum theory.
>
> C2: Granting (i)-(iii), the abstract-broad AB effect is necessary for understanding-with quantum theory.
>
> C3: The abstract-narrow AB effect is essential for understanding-what the concrete-narrow AB effect is (in the first place).
>
> C4: The abstract-broad AB effect is necessary for understanding-what the concrete-broad AB effect is (in the first place).

"Essential" in C1 and C3 is promoted to "necessary" in C2 and C4 when considering the abstract-broad AB effect. Here the full spectrum of structure of quantum mechanics only becomes evident when theorists make choices about how to formally implement I1-I3, thereby choosing a boundary condition (e.g., Dirichlet, Neumann, Robin) and deducing the relevant empirically confirmable physics in each case at hand.

Claims C1-C4 have straightforward implications for understanding-why the concrete AB effect occurs:

> C5: Requiring understanding-with quantum mechanics as a precondition for understanding-why the concrete-narrow AB effect occurs implies (via C1) that abstract-narrow AB effect is essential for understanding-why the concrete-narrow AB effect occurs; as a precondition for understanding-why the concrete-broad AB effect occurs implies (via C2) that the abstract-broad AB effect is necessary for understanding-why the concrete-broad AB effect occurs.
>
> C6: Requiring understanding-what the concrete-narrow AB effect is as a precondition for understanding-why the concrete-narrow AB effect occurs implies (via C3) that the abstract-narrow AB effect is essential for understanding-why the concrete-narrow AB effect occurs; as a precondition for understanding-why the concrete-broad AB effect occurs implies (via C4) that the abstract-broad AB effect is necessary for understanding-why the concrete-broad AB effect occurs.

Taking understanding-with and/or understanding-what as preconditions on understanding-why p does seem intuitive, e.g., as Newman (2017, 572) notes, both De Regt (2009, 2017) and Strevens (2013) "consider understanding a theory to be a necessary condition on understanding a scientific phenomenon." Moreover, any objection to the effect that such preconditions set the bar too high can be defused by noting that understanding admits of degrees (as is held by various philosophers of understanding cited above). In any case, I do not take a stance on the issue here.

Instead, let us consider worries to the effect that idealizations are neither necessary nor sufficient for understanding. First, starting with the non-sufficiency concern, I admit it is difficult to see how falsehoods can carry the explanatory-load in providing the explanation that affords understanding of why p occurs in the actual world. Nothing that I have said in this paper contradicts this basic intuition. It is facts about our world that carry the explanatory load in explaining why p occurs in the actual world—call this *veritism*. However, the AB effect case study suggest veritism isn't all there is to understanding. It shows how idealizations can play a *direct* role in understanding-with and understanding-what, and, consequently, an *indirect* role in understanding why p via understanding-with and understanding-what.

However, one may still worry. After all, many interlocutors in the debate on understanding, including factivists, agree that idealizations play *some* role in understanding. It is just that the role of idealizations is one of convenience, viz., ease of calculation, tractability, salience via flagging relevancies and irrelevancies. In reply, notice how understanding is afforded in the AB effect not in spite of idealizations but *in virtue* of them. Their primary role is not to ease calculation nor to flag relevancies/irrelevancies in the explanation of why the concrete AB effect occurs. It is to directly provide understanding-with and understanding-what, and perhaps indirectly afford a deeper sense of understanding-why.

Second, moving on the non-necessity concern, one may object: Surely a thorough understanding of the AB effect is exhibited if one understands all non-ideal cases. It might be useful to refer to the limiting case of no field outside the solenoid, but it can't be necessary for the conclusion that, in the non-ideal case, the magnetic field outside the solenoid is not responsible (or at least not solely responsible) for the effect. For instance, Sullivan & Khalifa (2019; 2, 15) recently argue that idealizations, which provide understanding of empirical phenomena, "merely alleviate the inconveniences required to use more accurate representations, either by simplifying calculations or by highlighting features that would otherwise be more difficult to spot" such that the "epistemic value that falsehoods have is not because of the understanding that they provide."

Although I understanding the intuition behind the "surely, … it can't be necessary" charge, I believe it is question-begging. As I reviewed above, the AB effect case study illustrates how idealizations are essential for understanding via claims C1-C6, wherein by "essential" I mean a strong pragmatic necessity that goes beyond issues of tractability. C2 and C4 upgrade such claims by championing necessity in the context of the broad AB effect. There one can provide an in-principle argument for why idealizations are necessary in terms of the formal implementation of I1-I3, viz., they allow for the emergence of a plethora of self-adjoint extension that "characterize all possible physical interaction of the particle with the solenoid border (sometimes obtained through non-trivial limit procedures)" (de Oliviera and Pereira 2010, 2). Such necessity claims have to do with understanding the structure of quantum mechanics (C2), understanding what the broad AB is in the first place (C4), and only indirectly with the manifestation of the effect in the laboratory (C5-C6).

Moreover, as noted above, the AB effect case study shows that understanding-with and understanding-what is afforded *in virtue* of non-actual but possible worlds according to quantum mechanics, which correspond to the abstract AB effect and are naturally construed as idealizations. If such worlds/scenarios count as "falsehoods," and gain in understanding-with and understanding-what counts as "epistemic value," then I'm in disagreement with Sullivan & Khalifa (2019). Having said that, Sullivan & Khalifa (2019, 3) are explicit about restricting they're "critique to positions arguing that idealizations have epistemic value in virtue of providing understanding of empirical phenomena…" This is not the main concern of my paper. C5-C6 state that idealizations are *indirectly* essential/necessary to understanding why the AB effect occurs only insofar as having understanding-with and understanding-what are preconditions on understanding-why. Sullivan & Khalifa (2019) can reject such preconditions, or claim that such preconditions only enhance one's understanding-why (since all notions of understanding discussed admit of degrees), or show how their "3D" approach deflates my account of understanding-with and understanding-what in the AB effect, or else reject the idea that understanding-with and understanding-what have epistemic value, or reject the idea that idealizations as I use them are "falsehoods." Any such development with would be interesting and welcome.

8. An Objection & Reply

One may worry that it isn't clear if I'm developing a new perspective of the AB effect using notions of understanding or articulating a new account of understanding using the AB effect. As an anonymous reviewer notes: If the former, how does the discussion of the AB effect go beyond the starting point of Earman (2019)? If the latter, what exactly is being added to the literature on scientific understanding?

In reply, I'm undertaking two complimentary tasks. On the one hand, I'm developing the idea that idealizations in the AB effect afford understanding by appealing to the concepts of understanding-with and understanding-what, and situating them in the understanding literature. Second, I'm noting that the manner by which idealizations afford understanding-with and understanding-what in the AB effect case study is importantly different from ways that such notions have been discussed. As for the former, although Earman's (2019) discussion is thorough and masterly, it doesn't develop how exactly idealizations afford understanding specifically, nor are any connection made to epistemological and philosophy of science literature on understanding. One is left wondering how and in what sense do the idealizations in the AB effect afford understanding. Appealing to the notions of understanding-with and understanding-what, while situating them in the literature, my analysis sheds light on such issues, adding the distinction between the abstract and concrete manifestation of the AB effect, and making connections with various contributions in the philosophy of understanding (e.g., Dellsén (2020), De Regt (2017), Kelp (2015), Khalifa (2017), Le Bihan (2017), Levy (2019), Newman (2017), Strevens (2008, 2013, 2017), as well as some of the philosophy of physics literature (e.g., Healey 1997, 1999; Maudlin 1998).

Still, it may be objected that since many readers take a very central aspect of understanding the AB effect to be a discussion of the causal-mechanistic explanation of the effect, and since this paper explicitly avoids such discussion, the contribution to our understanding of the AB effect is limited. However, and first, it isn't clear to me that a discussion of the causal-mechanistic explanation of the AB effect is necessary in order to contribute to our understanding of the effect. If such a discussion were necessary, then it would imply that works by, e.g., Batterman (2003), Earman (2019), Shech (2018), are not genuine contributions (which I take to be a reductio). It also begs the question against part of the upshot of these studies to the effect that the debate about the ontological causal-mechanistic explanation of the effect is "is largely a red herring" (Batterman 2003, 552).

Second, I explain above (in Section 5) that if one is partial to a causal-mechanistic account of explanation, and holds that (say) understanding-with is constituted by the relevant explanatory knowledge related to all phenomenon (concrete-empirical and abstract) that are within the purview of a theory, then such an account of understanding-with will likely attend to issues having to do with (ii) theoretical foundations and (iii) intertheoretic relations. For instance, if the AB effect is caused by a non-local and physically significant vector potential, then this implies that in quantum mechanics (as opposed to classical physics) vector potentials are real, field-like entities and that quantum theory is radically more indeterministic than previously thought. Surely knowing such information is part and parcel of understanding quantum theory and this is part of what I have suggested in this paper.

As for the latter, the important case study of the AB effect hasn't been discussed in the context of the scientific understanding literature. I have attempted to show how the notions of understanding that arises are in slight tension with some of the existing work. This has led me to argue that there are possible (but non-actual) worlds according to quantum mechanics, which are most naturally construed as idealizations, and are necessary/essential for understanding both quantum mechanics and what the AB effect is along the qualified lines of C1-C6. I have also suggested an outline of an account of understanding-with, and characterized a notion of understanding-what a phenomenon is in the first place, that I have not seen explicitly discussed in the literature. In order to further stress these points, I end my reply to this objection by comparing my claims with Potochnik (2017).

Potochnik's view is that science is in the business of representing causal patterns—recurrent regularities that are embodied in phenomena and manifest dependence relations—and that due to both the limited cognitive abilities of humans, and the "multifactorial miasma" and complexity of the world, such representations must be idealized (74). An idealization is a representation that represents a target as if it has properties that it does not have, for some purpose of interests such as explanation, prediction, etc. Given that idealizations are ubiquitous in science and, strictly speaking, false, it follows that science does not aim at truth. In its place, Potochnik holds that science aims at understanding in order to support human cognitive and practical ends.

I am sympathetic to aspects of Potochnick's view, especially the positive role that Potochnik identifies for idealizations in science, e.g., they promote understanding directly. I agree that, generally, "idealizations are used to set aside complicating factors to help scientists discern the causal patterns they are primarily interested in" (47). However, I submit that the primary role that idealizations play in the AB effect is importantly different in a number of ways. First, much of the target of study in the AB effect case study is the idealized, abstract AB effect itself. So here an idealization is *not* a representation that represents a target as if it has properties that it does not have. Second, as noted, the idealizations in the AB effect do not (only) simplify the corresponding experimental situation; they (also) complicate the situation through the emergence of formally distinct implementations of I1-I3 with diverse physical predictions.

Last, it is crucial to note that claims (C1 and C3) to the effect that abstract AB effect is "essential," and even more so that it is "necessary" (as in C2 and C4), for understanding-with and understanding-what ought not to be confused with considerations having to do with the complexity of the world and the limited cognitive abilities of humans. To be clear, on my account, even if humans had the calculating prowess of a Laplacian demon, they would still appeal to the abstract AB effect for understanding-with and understanding-what. Said differently, the positive role played by idealization in the AB effect case study does not stem from the world's complexity nor human's limited cognitive capacities.

9. Conclusion.

The history and philosophy involved in the AB effect case study is both fascinating and suggests that idealizations provide genuine understanding. But what is meant by "understanding" in this context? Drawing on understanding literature, and situation my own use of objectual understanding of theory and phenomenon in relation to said literature, I have argued that idealizations are directly essential/necessary for understanding-with quantum theory and understanding-what the AB effect is in the first place (C1-C4); with qualification, they are also indirectly essential/necessary for understanding-why the AB effect occurs (C5-C6). Along the way I noted that the AB effect case study, in interaction with recent literature (e.g. Newman (2017), Le Bihan (2017), and Levy (2019)), suggests an outline of an account of understanding-with. In brief:

> Subject S has understanding of a theory T to the extent that S is able to (i) draw accurate and general inferences about phenomena governed by T and navigate the modal space associated with T, (ii) answer various foundational question about T, and (iii) elucidate the relations of T to predecessor and competing theories and interpretations.

Future study includes developing said account, as well as exemplifying how the notions of understanding-with and understanding-what discussed here arise in other case studies.

Appendix.

The standard Hamiltonian operator for a charged particle in a magnetic field takes the form of $H = (\boldsymbol{p} - \frac{e}{c}\boldsymbol{A})^2$, where $\mathbf{A}$ is the vector potential, $\boldsymbol{p} = -i\nabla$ is the momentum operator, e and c are constants, and units have been chosen so that the mass of the electron is ½ and $\hbar = 1$. In the context of the AB effect, the idealization I1-I3 give rise to the following differential operator: $H^{\boldsymbol{A}_\infty} = (\boldsymbol{p} - \frac{e}{c}\boldsymbol{A}_\infty)^2$. In the Coulomb gauge of classical electromagnetism with polar coordinate (ρ, z, θ), $\rho \coloneqq (x^2 + y^2)^{1/2}$, and with the z-axis chosen as the axis of the infinite cylindrical solenoid S_∞ of radius R, $\boldsymbol{A}_\infty$ takes the following form:

$$(\boldsymbol{A}_\infty)_z = (\boldsymbol{A}_\infty)_\rho = 0$$

$$(\boldsymbol{A}_\infty)_\theta = \frac{\Phi_\infty}{2\pi\rho} \text{ for } \rho \geq R$$

$$(\boldsymbol{A}_\infty)_\theta = \frac{\Phi_\infty \rho}{2\pi R^2} \text{ for } 0 \leq \rho \leq R$$

Given I1-I3, the configuration space for electron is $\mathbb{R}^3 \setminus S_\infty$ since the probability of finding an electron in S_∞ is zero. The natural (initial) domain for $H^{\boldsymbol{A}_\infty}$ is $D(H^{\boldsymbol{A}_\infty}) = C_0^\infty(\mathbb{R}^3 \setminus S_\infty)$ which is dense in $\mathcal{H} = L^2(\mathbb{R}^3 \setminus S_\infty)$, where $C_0^\infty(\mathbb{R}^3 \setminus S_\infty)$ are the smooth functions of compact support on $\mathbb{R}^3 \setminus S_\infty$, and $L^2(\mathbb{R}^3 \setminus S_\infty,)$ is the Hilbert space of complex valued square integrable functions on $\mathbb{R}^3 \setminus S_\infty$. Stone's theorem (Reed and Simon 1980, 266-268, Theorem VIII.8), implies that the dynamics of a quantum system are determined by a unitary group, where the infinitesimal generator of such a group must be a self-adjoint operator or an essentially self-adjoint operator such that there is a unique extension to a larger domain on which the operator is self-adjoint. $H^{\boldsymbol{A}_\infty}$ is not essentially self-adjoint on its domain $D(H^{\boldsymbol{A}_\infty})$, but it has an (infinity-fold) infinity of self-adjoint extensions with corresponding boundary conditions on the wave function at the border of the solenoid (de Oliveira and Pereira 2010).

The assumption of impenetrability implies that the normal component $\boldsymbol{j}_N$ of the electron probability current $\boldsymbol{j}$ must vanish at the solenoid boundary (where $\boldsymbol{j} := -i(\psi^*\nabla\psi - \psi\nabla\psi^*)$ (Earman 2019, 1999-2000). There are several sufficient (but not necessary) conditions for implemented the requirement using different boundary conditions including the Dirichlet boundary conditions ($\psi = 0$) that Aharonov and Bohm (1959) used, Neumann boundary conditions ($\nabla\psi = 0$), or Robin boundary conditions ($\nabla\psi = r\psi, r \in \mathbb{R}$) (de Oliveira and Pereira 2010, 7-8). Importantly, different boundary conditions with essentially self-adjoint extension of $H^{\boldsymbol{A}_\infty}$ correspond to distinct dynamical evolutions with diverse empirical predictions for scattering experiments (Fig. 3). de Oliveira and Pereira (2010, 28) hold that this confirms "the presence of the AB effect … in different self-adjoint extensions," which I refer to in the main text as the *AB effect broadly construed*.

The *AB effect narrowly* construed concerns an interference pattern in a double-slit experiment that is shifted by an amount $\Delta x = \frac{l\lambda e}{2\pi d\hbar}\Phi_\infty$, where λ is the wave length of the electron, e is the charge of the electron, $\hbar$ the reduced Planck constant, d the distance between slits, l the distance between the double-slit screen and detector screen, and Φ_∞ is the magnetic flux through S_∞ when the solenoid is turned on. Making use of $\bar{H}_{AB}^{\boldsymbol{A}_\infty}$, which is a self-adjoint extension of $H^{\boldsymbol{A}_\infty}$ corresponding to Dirichlet boundary conditions, Aharonov and Bohm (1959) provide a semi-classical approximation of the relative phase shift by assuming that the exact solution ψ_1 (of the Schrödinger equation with $\bar{H}_{AB}^{\boldsymbol{A}_\infty}$) when the solenoid is turned *on* is well approximated by the solution $\psi_1{}'$ when the solenoid is turned *off* multiplied by a phase factor $e^{i\frac{e}{c}\Lambda_1}$, where $\Lambda_1 = \int_1 \boldsymbol{A} \cdot d\boldsymbol{r}$ is a scalar function corresponding to a line integral taken along a path in a simply connected region around the solenoid (and the same assumption is made for $\psi_2 = \psi_2{}' e^{i\frac{e}{c}\Lambda_2}$) (see Earman (2019, Section 5.1) for details). This assumption is dubbed the Aharonov-Bohm *Ansatz* (Ballesteros and Weder 2009, 2011; Earman 2019), and it facilitates the prediction of a relative phase change $e^{i\theta} = e^{i\frac{e}{c}\Phi_\infty}$, where the the shift in interference patters depends on the magnetic flux Φ_∞. Moreover, Ballesteros and Weder (2009, 2011) show that for high velocity Gaussian electron wave packets the AB Ansatz is a good approximation for the type of systems used to confirm the AB effect experimentally (viz., with toroidal solenoids), e.g., Tonomura *et al*. (1982, 1986).

References

Aharonov, Y., D. Bohm. 1959."Significance of electromagnetic potentials in the quantum theory." *Physical Review* 115: 485-91.

Bain, J. 2016. Emergence and the mechanism in the fractional quantum Hall effect. *Studies in History and Philosophy of Modern Physics*, 56, 27–38.

Ballesteros, M., R. Weder. 2009. "The Aharonov–Bohm effect and Tonomura et al. experiments: Rigorous results." *Journal of Mathematical Physics* 50: 122108.

Ballesteros, M., R. Weder. 2011. "Aharonov–Bohm effect and high-velocity estimates of solutions to the Schrodinger equation." *Communications in Mathematical Physics* 303(1): 175-211.

Batterman, R. 2003. "Falling cats, parallel parking, and polarized light." *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 34: 527–557.

Baumberger, C. 2019. "Explicating Objectual Understanding: Taking Degrees Seriously." *Journal for General Philosophy of Science* 50 (3):367-388. doi: 10.1007/s10838-019-09474-6.

Baumberger, C., Beisbart, C., and G. Brun. 2017. "What is Understanding? An Overview of Recent Debates in Epistemology and Philosophy of Science." in Grimm *et al*. pp. 1-34.

Belot, G. 1998. "Understanding Electromagnetism." *British Journal for Philosophy of Science* 49(4): 531–555.

Bocchieri, P., & Loinger, A. 1978. Nonexistence of the Aharonov–Bohm effect. Nuovo Cimento, 47A, 475–482.

Bocchieri, P., & Loinger, A. 1981. Charges in multiply connected spaces. Nuovo Cimento, 66, 164–172.

Bohm, D., & Hiley, B. J. 1979. On the Aharonov–Bohm Effect. Nuovo Cimento, 52A, 295–307.

Bokulich, A. 2008. *Reexamining the Quantum-Classical Relation: Beyond Reductionism and Pluralism*. Cambridge: Cambridge University Press.

Boyer, T. H. 2000a. Does the Aharonov-Bohm Effect Exist?. *Foundations of Physics*, 30(6), 893-.905.

Boyer, T. H. 2000b. Classical Electromagnetism and the Aharonov-Bohm Phase Shift. *Foundations of Physics*, 30(6), 907-932.

Boyer, T. H. 2006. Darwin-Lagrangian analysis for the interaction of a point charge and a magnet: Considerations related to the controversy regarding the Aharonov-Bohm and Aharonov-Casher phase shifts. *Journal of Physics* A, 39, 3455-3477.

Boyer, T. H. 2008. "Comment on the Experiments Related to the Aharonov-Bohm Phase Shift." *Foundations of Physics* 38: 398-505.

Brun, G. and Baumberger, C. 2017. 'Dimensions of Objectual Understanding', in Grimm *et al*. pp. 165–89.

Bunge, M. 2015. Does the Aharonov–Bohm Effect Occur? *Foundations of Science*. 20:129–133.

Carter, J. A. and Gordon, E. C. 2016. 'Objectual Understanding, Factivity, and Belief', in M. Grajner and P. Schmechtig (eds), Epistemic Reasons, Norms and Goals, Berlin: De Gruyter, pp. 423–42.

Chambers, R. G. 1960. "Shift of an Electron Interference Pattern by Enclosed Magnetic Flux." Physical Review Letter 5(1): 3-5.

Crease, R. P. 2002. *Physics World* 15(9): 19-20.

Daugherty, J. 2021. "The non-ideal theory of the Aharonov–Bohm effect." *Synthese*. 198:12195–12221.

Dellsén, F. 2020. "Beyond Explanation: Understanding as Dependency Modeling." *British Journal for the Philosophy of Science*. 71:1261–1286.

de Oliveira, C. R., M. Pereira. 2008. "Mathematical Justification of the Aharonov-Bohm Hamiltonian." Journal of Statistical Physics 133:1175-1184.

de Oliveira, C. R., M. Pereira. 2010. "Scattering and Self-adjoint extensions of the Aharonov-Bohm Hamiltonian." *Journal of Physics A: Mathematical and Theoretical* **43:**1-29.

de Oliveira, C. R., M. Pereira. 2011. "Impenetrability of Aharonov-Bohm Solenoids: Proof of Norm Resolvent Convergence." *Letters in Mathematical Physics* 95: 41-51.

De Regt, H. W. 2009. "Understanding and Scientific Explanation," in De Regt *et al*. pp. 21–42.

De Regt, H. W. 2017. *Understanding Scientific Understanding*, Oxford University Press.

De Regt, H. W., Forthcoming a. "Can Scientific Understanding be Reduced to Knowledge?" in Khalifa, K., Lawler, I. and E. Shech. (eds.) Forthcoming.

De Regt, H. W., Forthcoming b. "Frenemies or friends? A Reply to Kareem Khalifa" in Khalifa, K., Lawler, I. and E. Shech. (eds.) Forthcoming.

De Regt, H. W., Gijsbers, V. 2017. "How False Theories Can Yield Genuine Understanding." in Grimm *et al*. 2017. pp. 50-75.

De Regt, H. W., Leonelli, S. and K. Eigner (eds.) 2009. *Scientific Understanding,* Pittsburgh: University of Pittsburgh Press.

Earman, J. 2019. "The Role of Idealization in the Aharonov-Bohm Effect." *Synthese* 196:1991–2019.

Ehrenberg, W., R. W. Siday, R. W. 1949. "The refractive index in electron optics and the principles of dynamics." *Proceedings of the Physical Society London: Section B* 62(1): 8-21.

Elgin, C. 2017. "Exemplification in Understanding." in Grimm *et al*. 2017. pp. 76-92.

Grimm, S. R., Baumberger, C., and S. Ammon. 2017. *Explaining Understanding: New Perspective from Epistemology and Philosophy of Science*. New York, NY: Routledge.

Healey, R. 1997. "Nonlocality and the Aharonov–Bohm effect." *Philosophy of Science*, 64, 18–41.

Healey, R. 1999. "Quantum Analogies: a Reply to Maudlin." *Philosophy of Science*, 66, pp. 440-447.

Healey, R. 2007. *Gauging What's Real: The Conceptual Foundations of Contemporary Gauge Theories.* Oxford University Press: New York.

Hempel, C. and P. Oppenheim., 1948, 'Studies in the Logic of Explanation.', Philosophy of Science, 15: 135–175. Reprinted in Hempel, 245–290, 1965a.

Jösson, C. 1961. *Zeitschrift für Physik* 161, 454-474.

Kelp, Christoph. 2015. "Understanding phenomena." *Synthese* 192 (12):3799–3816. doi: 10.1007/s11229-014-0616-x.

Klein, U. 1979. The inadmissibility of non-stokesian vector potentials in quantum theory. Lettere al Nuovo Cimento, 25, 33–37.

Khalifa, K. 2011. "Understanding, knowledge, and scientific antirealism," Grazer Philosophische Studien 83: 93-112.

Khalifa, K. 2013. "Is understanding explanatory or objectual?" *Synthese* 190: 1153-1171.

Khalifa, K. 2017. *Understanding, Explanation, and Scientific Knowledge*, Cambridge University Press.

Khalifa, K. Forthcoming a. "Should Friends and Frenemies of Understanding be Friends? Discussing de Regt" in Khalifa, K., Lawler, I. and E. Shech. (eds.) Forthcoming.

Khalifa, K. Forthcoming b. "Onwards, My Friend! Reply to De Regt" in Khalifa, K., Lawler, I. and E. Shech. (eds.) Forthcoming.

Khalifa, K., Lawler, I. and E. Shech. (eds.) Forthcoming. *Scientific Understanding and Representation Modeling in the Physical Sciences*. Routledge.

Ladyman, J. 2008. "Idealization" *The Routledge Companion to Philosophy of Science*. Eds. Psillos, S. and Curd, M. London and New York: Routledge Taylor & Francis Group: 358-366.

Le Bihan, S. 2017. "Enlightening Falsehoods: A Modal View of Scientific Understanding." In Grimm *et al*. 2017. pp. 111-136.

Levy, A. 2019. "Metaphor and Scientific Explanation." In Godfrey-Smith, P. and Levy, A. *The Scientific Imagination*. Oxford: Oxford University Press.

Lyre, H. 2009. "Aharonov–Bohm effect." In D. Greengerger,C.Hentschel,&F.Weinert (Eds.), *Compendium of quantum mechanics* (pp. 1–3). Berlin: Springer.

Pleijel, H. 1937. Award ceremony speech. NobelPrize.org. Nobel Media AB 2020. Thu. 13 Aug 2020. <https://www.nobelprize.org/prizes/physics/1937/ceremony-speech/>

Popper, K.R. 1972. *Objective Knowledge: An Evolutionary Approach*, Oxford: Clarendon Press.

Malament, D. B. 2008. "Norton's Slippery Slope." *Philosophy of Science*,75, 799-816.

Mattingly, J. 2006. "Classical fields and quantum time-evolution in the Aharonov–Bohm effect." *Studies in History and Philosophy Science Part B: History and Philosophy of Modern Physics*, 37, 243–262.

Maudlin, T. 1998. “Healey on the Aharonov–Bohm effect.” *Philosophy of Science*., 65, 361–368.

Möllenstedt, G., W. Bayh. 1962. „Kontinuierliche Phasenschiebung von Elektronenwellen im kraftfeldfreien Raum durch das magnetische Vektorpotential eines Solenoids.“ *Zeitschrift für Physik* 169:263.

Newman, Mark P. 2017. "Theoretical Understanding in Science." *The British Journal for the Philosophy of Science* 68 (2):571-595. doi: 10.1093/bjps/axv041.

Norton, J. 2008. The dome: an unexpectedly simple failure of determinism. *Philosophy of Science*, 75, 786-98.

Norton, J. 2012. Approximations and idealizations: Why the difference matters. *Philosophy of Science*, 79, 207–232.

Shech, E. 2015. “Two Approaches to Fractional Statistics in the Quantum Hall Effect: Idealizations and the Curious Case of the Anyon” Foundations of Physics, Volume 45, Issue 9: 1063-110.

Shech, E. 2018. “Idealizations, Essential Self-Adjointness, and Minimal Model Explanation in the Aharonov-Bohm Effect.” *Synthese*, 195:4839–4863.

Shech, E. 2019. “Philosophical Issues Concerning Phase Transitions and Anyons: Emergence, Reduction, and Explanatory Fictions.” *Erkenntnis*, (84)3:585–615.

Shech, E. and A. Gelfert. 2019. “The Exploratory Role of Models and Idealizations.” *Studia Metodologiczne – Dissertationes Methodologicae*. ISSN 0039-324X *Issue on Culture(s) of Modelling in Science(s)* (39)

Shech, E. and P. McGivern. 2021. “Fundamentality, Scale, and the Fractional Quantum Hall Effect.” *Erkenntnis*, 86:1411–1430.

Strevens, M. 2008. *Depth: An Account of Scientific Explanation,* Cambridge: Harvard University Press.

Strevens, M. 2013. “No understanding without explanation,” *Studies in History and Philosophy of Science* 44: 510-515.

Strevens, M. 2017. “How Idealizations Provide Understanding.” in Grimm *et al*. 2017. pp. 37-49.

Strocchi, F., &Wightman, A. S. (1974). Proof of the charge superselection rule in local relativistic quantum field theory. Journal of Mathematical Physics, 15, 2189–2224.

Stuart, M. T. 2018. “How Thought Experiments Increase Understanding.” in Michael T. Stuart, Yiftach J. H. Fehige & James Robert Brown (eds.), The Routledge Companion to Thought Experiments. London: Routledge. pp. 526-544.

Tonomura, A. et al. 1982. “Observation of the Aharonov-Bohm Effect by Electron Holography.” *Physical Review Letters* 48:1443.

Tonomura, A. et al. 1986. “Evidence for Aharonov-Bohm Effect with Magnetic Field Completely Shielded from Electron Wave.” Physical Review Letter 56:792-795.

Tonomura, A. 1999. *Electron Holography*. Berlin: Springer-Verlag.

Vaidman, L. 2012. Role of potentials in the Aharonov-Bohm effect. Physical Review A 86, 040101

Valente, G. 2019. “On the paradox of reversible processes in thermodynamics.” *Synthese*. 196:1761–1781.

Wallace, D. 2014. Deflating the AB effect. https://arxiv.org/pdf/1407.5073.pdf

Wang, R. F. 2015. Absence of electric Aharonov-Bohm effect due to induced charges. *Scientific Reports*. doi:10.1038/srep14279.

Wilczek, F. 1982. ‘Quantum Mechanics of Fractional-Spin Particles’, *Physical Review Letters*, 49, pp. 957-959.

Wilczek, F. (ed.) 1990. *Fractional Statistics and Anyon Superconductivity*. Singapore: World Scientific.