The Trouble with Weak Values

Jacob A. Barandes*

February 9, 2026

Abstract

In quantum theory, a weak value is a complex number with a somewhat technical definition: it is a ratio whose numerator is the matrix element of a self-adjoint operator and whose denominator is the inner product of a corresponding pair of state vectors. Weak values first appeared in the research literature in a pair of papers in 1987 and 1988, and were originally defined as the results of a special kind of experimental protocol involving non-disturbing measurements combined with an explicit form of post-selection. In the years since, subsequent papers on weak values have produced a number of important practical spin-offs, including new methods for signal amplification and quantum-state tomography. The present work is not concerned with those practical spin-offs, but with historical and ongoing attempts to assign weak values a transparent, single-system interpretation, as well as efforts that invoke weak values to make a number of exotic claims about the properties and behavior of individual quantum systems. This paper challenges these interpretational claims by arguing that they involve several forms of fallacious reasoning.

"It is the sculptor's power, so often alluded to, of finding the perfect form and features of a goddess, in the shapeless block of marble; and his ability to chip off all extraneous matter, and let the divine excellence stand forth for itself." —Robert Allyn, 1858

1 Introduction

In quantum theory, a 'weak value' has a rather formal definition: it is a complex-valued numerical ratio whose numerator consists of the matrix element of a self-adjoint operator and whose denominator consists of the Hilbert-space inner product of a corresponding pair of state vectors. Many formal constructions in quantum theory admit clear physical interpretations. To what extent is such an interpretation available for weak values? Could an interpretation of this kind reveal underlying facts about quantum systems that would otherwise be empirically inaccessible?

*Departments of Philosophy and Physics, Harvard University, Cambridge, MA 02138; jacob_barandes@harvard.edu; ORCID: 0000-0002-3740-4418

The origins of weak values go back to a 1987 paper by Aharonov, Albert, Casher, and Vaidman (AACV) (1987), together with a 1988 paper by Aharonov, Albert, and Vaidman (AAV) (1988). According to the website of the *Physical Review*, the AAV paper alone has received over 2,200 citations, and there are many thousands more papers in the research literature on weak values today.

The discovery of weak values was a landmark and deservedly celebrated development in science. In particular, experimental methods for extracting weak values have led to many important practical applications, including new methods for amplifying detector signals, estimating small evolution parameters, and quantum-state tomography (Dressel et al. 2014).

The present work will not aim its critical analysis at those practical applications. Instead, this paper will focus on interpretational claims about weak values in the research literature, as well as on ways in which weak values have been used to justify still other interpretational claims about specific kinds of quantum systems.

One sees these sorts of interpretational claims in the original AACV and AAV papers, which introduced an ensemble-based protocol for indirectly extracting weak values experimentally, with post-selection playing a central role. These two papers were partly inspired by a 1964 paper by Aharonov, Bergmann, and Lebowitz (ABL) (1964), which studied an interesting kind of conditional-probability formula, called the 'ABL rule,' that was defined in terms 'pre-selected' and 'post-selected' quantum states.[1]

These interpretational claims have generated much controversy over the years. A prominent example was a flurry of arguments and counterarguments arising from a critical paper by Ferrie and Combes that was published in *Physical Review Letters* in 2014 and centered on, among other things, questions over whether weak values were truly quantum-mechanical (Ferrie, Combes 2014; Vaidman 2014; Cohen 2014; Aharonov, Rohrlich 2014; Brodutch 2015; Ferrie, Combes 2015; Romito, Jordan, Aharonov, Gefen 2016).

The AACV paper introduced an explicit example that the AAV paper later developed into a more general framework. The present work will therefore focus primarily on the AAV paper.

Section 2 will review several preliminary topics, including relevant fallacies of reasoning and the textbook Dirac-von Neumann axioms of quantum theory. Section 3 will start with a first look at weak values, and then proceed to lay out the standard experimental protocol for extracting them, followed by a discussion of their interpretation and history. Section 4 will present a critical analysis of two foundational applications of weak values—to "quantum Cheshire Cats" and to Bohmian trajectories. Section 5 will conclude with a brief summary and a look at larger issues.

[1] Although both the AACV and AAV papers used pre-selection and post-selection in essential ways, only the AACV paper mentioned the ABL rule explicitly, and then only in passing. The AACV and AAV papers therefore logically stood on their own. In particular, to whatever extent the discovery of weak values has been a success, that success does not provide retroactive evidentiary support for the validity of the ABL rule (Barandes 2026).

2 Preliminaries

2.1 Three Relevant Fallacies of Reasoning

The critical analysis ahead will make explicit and repeated reference to three types of fallacious reasoning introduced in other work (Barandes 2026): the ensemble fallacy, the post-selection fallacy, and the measurementist fallacy. It will be worth reviewing these fallacies briefly here.

In physics, one sometimes studies a single object or system, and sometimes one instead imagines or physically constructs a large number, or ensemble, of systems that share some important collection of key characteristics.

An observable feature that one can operationally assign to a single system will, appropriately, be called a single-system observable. An example might be the mass of a star, or the spectrum of possible measurement outcomes belonging to some observable of a quantum system. By contrast, an observable feature that only makes sense at the irreducibly collective or emergent level of an ensemble as a whole will be called an ensemble observable.

Although it is sometimes convenient to study a single-system observable by making use of an ensemble, such a method does not collapse the categorical distinction between single-system observables and ensemble observables (Hance, Rarity, Ladyman 2023, Section VI). Collapsing that categorical distinction constitutes a form of fallacious reasoning:

> The Ensemble Fallacy: The category error of attempting to identify a single-system observable with an ensemble observable, or to draw direct inferences about a single-system observable from an ensemble observable without independent rigorous justification. (1)

In statistics, one needs to be very careful to ensure that the way in which one implicitly or explicitly culls an ensemble down to a subensemble does not introduce spurious correlations and other statistical artifacts, called selection biases (Hernán, Hernández-Díaz, Robins 2004). A famous example is Berkson's paradox (Berkson 1947), or collider bias (Cole et al. 2009), which refers to problems that arise if one is insufficiently careful about conditioning an ensemble by a form of after-the-fact selection, or post-selection. Berkson, in particular, noticed that when looking at patients in a hospital, certain diseases were negatively correlated, but this negative correlation was a form of selection bias, because if a patient did not have the symptoms of one disease, then the patient would need to have symptoms of a different disease in order to be admitted to the hospital in the first place.

Yet, rather than trying to avoid or control for accidental post-selection, some papers in the quantum-foundations research literature actually *embrace* post-selection, and then ascribe their exotic-looking statistical results to the mysterious nature of quantum theory itself. However, the mere fact that one is working with quantum systems does not give one license to make the sorts of inferences in the midst of post-selection that would otherwise be suspect in conventional statistics. It is worth giving this sort of widespread error a name:

The Post-Selection Fallacy: Drawing erroneous conclusions about a system or ensemble due to the implicit or explicit use of post-selection. (2)

When all else fails, it can be tempting to declare that a favored interpretation of a theoretical quantity holds, or that work aimed at developing that theoretical quantity is worthwhile, merely because that theoretical quantity can be measured experimentally in the laboratory. This view is itself a fallacious form of reasoning, and merits having its own name:

The Measurementist Fallacy (or Measurementism): If a quantity of ambiguous interpretation can be measured experimentally, then experiments alone can provide confirmatory support for a favored interpretation of that quantity, or a justification for theoretical work that led to the consideration of that quantity. (3)

2.2 The Dirac-von Neumann Axioms

To establish the ground rules for the discussion ahead, it will be prudent to lay out the Dirac-von Neumann (DvN) axioms (Dirac 1930, von Neumann 1932), which underlie orthodox or textbook quantum theory:

1. Quantum states: Each quantum system is assigned a Hilbert space of some finite or infinite dimension. In the most general case, the system's quantum state is a self-adjoint, positive semidefinite operator with unit trace, called a density matrix, or density operator, ρ:

$$\rho = \rho^\dagger \geq 0, \quad \mathrm{tr}\,\rho = 1. \tag{4}$$

If the system's density matrix has rank greater than one, then it is said to be mixed. In the special case in which the system's density matrix is rank-one, or pure, it can be factorized as the outer product of a unit-norm vector $|\Psi\rangle$ with its adjoint $\langle\Psi| \equiv |\Psi\rangle^\dagger$:

$$\rho = |\Psi\rangle\langle\Psi|, \quad \langle\Psi|\Psi\rangle = 1 \quad [\text{if } \rho \text{ is rank-one}]. \tag{5}$$

This vector, called a state vector, or wave function, is uniquely defined up to an irrelevant overall phase factor:

$$|\Psi\rangle \cong e^{i\theta}|\Psi\rangle. \tag{6}$$

Given multiple quantum systems that are subsystems of a composite system, their individual Hilbert spaces $\mathcal{H}_1, \ldots, \mathcal{H}_n$ combine as a tensor product to define the Hilbert space of the composite system:

$$\mathcal{H}_{\text{composite}} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n. \tag{7}$$

The density matrices of the subsystems are related to the density matrix of the composite system via the partial-trace operation:

$$\rho_1 = \mathrm{tr}_{\mathrm{not}\ 1}[\rho], \quad \ldots, \quad \rho_n = \mathrm{tr}_{\mathrm{not}\ n}[\rho]. \tag{8}$$

2. Unitary time evolution: The quantum state of a closed quantum system, meaning a quantum system that is not engaged in mutual interactions or information exchange with any other quantum systems, changes with time according to a time-indexed family of unitary transformations that make up the system's time-evolution operator $U(t \leftarrow t_0)$:

$$\rho(t) = U(t \leftarrow t_0)\rho(t_0)U^\dagger(t \leftarrow t_0), \quad |\Psi(t)\rangle = U(t \leftarrow t_0)|\Psi(t_0)\rangle. \tag{9}$$

If the time-evolution operator is sufficiently smooth as a function of time, then one can define the Hamiltonian as the self-adjoint operator

$$H(t) \equiv i\hbar \frac{\partial U(t \leftarrow t_0)}{\partial t} U^\dagger(t \leftarrow t_0) = H^\dagger(t), \tag{10}$$

where $\hbar$ is the reduced Planck constant. One can then express the time evolution of the system's quantum state as a first-order differential equation—the von Neumann equation for density matrices or the Schrödinger equation for state vectors:

$$i\hbar \frac{\partial \rho(t)}{\partial t} = [H(t), \rho(t)], \quad i\hbar \frac{\partial |\Psi(t)\rangle}{\partial t} = H(t)|\Psi(t)\rangle. \tag{11}$$

3. Observables: A single quantum system has an associated collection of observables, each of which is represented by a self-adjoint operator:

$$A = A^\dagger. \tag{12}$$

If a measurement is carried out on an observable A, then the possible numerical outcomes are given by the operator's eigenvalue spectrum $\sigma(A)$.

4. The Born rule: Given a self-adjoint operator A representing one of the quantum system's observables, let P_a denote a self-adjoint, idempotent projection operator or eigenprojector associated with an eigenvalue a in the spectrum $\sigma(A)$:[2]

$$P_a = P_a^\dagger = P_a^2, \quad AP_a = aP_a. \tag{13}$$

These eigenprojectors form a projection-valued measure, or PVM, meaning that they satisfy the conditions of mutual exclusivity,

$$P_a P_{a'} = \delta_{aa'} P_a, \tag{14}$$

[2]For mathematical simplicity, the spectrum $\sigma(A)$ will be assumed here to be countable.

where $\delta_{aa'}$ is the usual Kronecker delta,

$$\delta_{aa'} = \begin{cases} 1 & \text{for } a = a', \\ 0 & \text{for } a \neq a', \end{cases} \tag{15}$$

as well as a completeness relation, or resolution of the identity,

$$\sum_a P_a = \mathbb{1}, \tag{16}$$

where $\mathbb{1}$ is the identity operator on the quantum system's Hilbert space $\mathcal{H}$. Then the probability $p(a)$ with which a measurement of A will yield the eigenvalue a is given by the Born rule:

$$p(a) = \mathrm{tr}(P_a \rho). \tag{17}$$

If the system's density matrix is rank-one, so that one can instead work with a state vector $|\Psi\rangle$, then the Born rule takes the form

$$p(a) = \mathrm{tr}(P_a |\Psi\rangle\langle\Psi|) = \langle\Psi|P_a|\Psi\rangle. \tag{18}$$

If the eigenprojector P_a is likewise rank-one, so that it admits a factorization of the form

$$P_a = |a\rangle\langle a|, \tag{19}$$

where $|a\rangle$ is an eigenvector of A with eigenvalue a,

$$A|a\rangle = a|a\rangle, \tag{20}$$

then the Born rule takes the simpler form

$$p(a) = |\langle a|\Psi\rangle|^2. \tag{21}$$

Given the Born rule, it follows that the expectation value $\langle A\rangle$ of an observable, meaning the statistical average of its possible numerical outcomes a weighted by their corresponding measurement-outcome probabilities $p(a)$,

$$\langle A\rangle \equiv \sum_a a\, p(a), \tag{22}$$

can be calculated from a density matrix ρ according to

$$\langle A\rangle = \mathrm{tr}(A\rho), \tag{23}$$

and can be calculated from a state vector $|\Psi\rangle$ according to

$$\langle A \rangle = \langle \Psi | A | \Psi \rangle. \tag{24}$$

5. Collapse: At the end of a measurement that yields an eigenvalue a corresponding to an eigenprojector P_a of an observable A, the system's quantum state reduces or collapses via a projection by P_a, followed by a renormalization to restore the trace or norm back to 1:

$$\rho \mapsto \frac{P_a \rho P_a}{\mathrm{tr}(P_a \rho)}, \qquad |\Psi\rangle \mapsto \frac{P_a|\Psi\rangle}{\sqrt{\langle \Psi | P_a | \Psi \rangle}}. \tag{25}$$

One can generalize these axioms to accommodate positive-operator-valued measures, or POVMs, but that generalization will not be needed in this paper.

According to the DvN axioms, only *measurements* are capable of triggering the third, fourth, and fifth axioms above, and only the fifth axiom takes the system to a final quantum state that *singles out* a specific measurement outcome. However, the DvN axioms famously do not define what a measurement is, or what kinds of systems can carry out measurements, so the DvN axioms are manifestly incomplete. This specific manifestation of incompleteness is called the measurement problem.[3]

A conceptually distinct issue is that even if one had a rigorous definition of which processes counted as measurements and which did not, there would still ostensibly be a gap between the narrow category of measurement processes and the seemingly much larger category of non-measurement phenomena that we think are happening around us all the time. This discrepancy is called the category problem in other work (Barandes 2025).

One cannot brush aside the category problem merely by appealing to the probabilities and expectation values provided by the DvN axioms,[4] because those are *measurement-outcome* probabilities and *measurement-outcome* expectation values, where the latter are numerical measurement outcomes statistically weighted by measurement-outcome probabilities. These are not probabilities or expectation values of phenomena just *happening* or *existing* in a broader categorical sense. Moreover, unless one is willing to engage in serious foundational work, one cannot choose to re-interpret all the measurement-outcome probabilities of the DvN axioms as referring to phenomena happening or existing, due to several no-go theorems, including the Kochen-Specker theorem (Bell 1966; Kochen, Specker 1967).

This paper's use of orthodox or textbook quantum theory, based on the DvN axioms, should not be taken as an endorsement of the textbook theory or the DvN axioms. The measurement problem and the category problem make clear that the textbook theory is likely incomplete, and there are

[3]Claims arise from time to time that decoherence alone is capable of solving the measurement problem. By itself, however, decoherence does not have have the resources to single out a unique measurement outcome, let alone with a specific probability, so further interpretational steps that go beyond the DvN axioms are necessary (Bacciagaluppi 2025). Due to space limitations, this paper will not say anything more about debates over the measurement problem.

[4]One finds just such an appeal in a popular textbook on quantum mechanics by Shankar (1994, Chapter 6, "The Classical Limit").

several candidate reformulations or 'interpretations' of quantum theory that seek to replace the DvN axioms with a more internally consistent set of axioms. However, these alternative reformulations generally make the same empirical predictions as the textbook theory for the sorts of experimental protocols that are related to weak values, so the textbook theory will be sufficient for the purposes of this paper.

Of course, there may be some new axiomatic framework for quantum theory on which the critical analysis presented by this paper no longer succeeds. It would be exciting to see the development of such a framework. This paper's overall argument is that unless or until such a new framework is found, it is not possible to sustain certain widely held interpretative claims about weak values.

It will be important to remember that the DvN axioms assign observables and their possible measurement outcomes at the level of a *single* quantum system. Given a single quantum system and a particular observable, one is justified in talking about which outcomes a measurement can yield for that single system, collectively forming the spectrum of that observable. Immediately after obtaining one such measurement outcome, one then applies the collapse axiom to that single system, thereby locking in that measurement outcome.

Importantly, the collapse axiom ensures that if the same observable is quickly measured again, by *any other* measuring device or observer, then the measurement outcome will be the same as before, within small error bars. This key consequence of the collapse axiom strongly motivates associating the given observable with that single system, at least in an empirical and operational sense.

In addition, notice that the measurement-outcome probabilities and expectation values provided by the DvN axioms are fully defined once one specifies a single system's quantum state together with the relevant self-adjoint operators belonging to that single system. Furthermore, these probabilities and expectation values *refer* to single-system measurement outcomes. However, the DvN probabilities and expectation values *themselves* can only be measured by considering large ensembles of systems. Hence, these probabilities and expectation values *themselves* are not single-system observables, in the strict empirical or operational sense used in this paper, and to treat them otherwise would be to commit the ensemble fallacy (1).[5]

In particular, if one encodes some sort of pattern into the probability of a given system, then that pattern will not be a single-system observable, but an ensemble observable. This observation will turn out to be important for the case of weak values, which are an abstract generalization of expectation values and not a generalization of single-system observables.[6]

[5]The AAV paper argues that in the limit in which a measuring device's pointer variable has a very large variance, the resulting measurement-outcome probability distribution for the observable to be measured will be very slightly peaked at the expectation value of that observable. However, this expectation value is then a property of a probability distribution, and, further, arises emergently from the interaction of the measuring device with the subject system as a unified whole, rather than being a single-system observable of the subject system alone. More importantly, as the AAV paper immediately points out, "One measurement like this will give no information," and then explains that this expectation value can only be measured for empirical and operational purposes by using a physical ensemble, so it is manifestly an ensemble observable.

[6]In the philosophy of probability, there are views, such as subjective or propensity theories (Hájek 2023), on which one can try to assign probabilities a metaphysical meaning as *properties* attached to a single system. However, empirically or operationally speaking, one can only measure probabilities at the level of an ensemble, so according to

3 Weak Values

3.1 A First Look at Weak Values

Given a quantum system with an observable represented by a self-adjoint operator $A = A^\dagger$ on its Hilbert space $\mathcal{H}$, together with a pair of state vectors $|\Psi\rangle, |\Psi'\rangle$ in its Hilbert space that have nonzero inner product $\langle\Psi'|\Psi\rangle \neq 0$, the corresponding 'weak value' A_w is formally defined as the complex-valued ratio

$$A_w \equiv A_w\big(\Psi', \Psi\big) \equiv \frac{\langle\Psi'|A|\Psi\rangle}{\langle\Psi'|\Psi\rangle} \quad \big[\text{assuming } \langle\Psi'|\Psi\rangle \neq 0\big]. \tag{26}$$

If $|\Psi'\rangle = |\Psi\rangle$ (up to an irrelevant phase factor), then the unit-norm condition $\langle\Psi|\Psi\rangle = 1$ from (5), together with the usual formula for the expectation value $\langle A\rangle = \langle\Psi|A|\Psi\rangle$ of an observable from (24), implies that the weak value reduces to the real-valued expectation value,

$$A_w(\Psi, \Psi) = \langle A\rangle, \tag{27}$$

as noted by Vaidman (1996a).[7] In the even more special case in which $|\Psi'\rangle = |\Psi\rangle = |a\rangle$ is an eigenvector of A (again up to irrelevant phase factors), the weak value reduces to the numerical measurement result a,

$$A_w(a, a) = a, \tag{28}$$

which is again real-valued.

There is nothing about the definition (26) of weak values, or about the special cases (27) and (28), that would immediately explain why they are called 'weak values.' The reason for the 'weak' modifier in their name is that the standard experimental protocol for extracting weak values requires special kinds of interactions between subject systems and measuring devices. These special kinds of interactions are called 'weak measurements' because they only slightly perturb the systems that are involved. Weak measurements will be reviewed in Subsection 3.3, and were originally introduced in both the 1987 AACV paper (Aharonov, Albert, Casher, Vaidman 1987) and in the 1988 AAV paper (Aharonov, Albert, Vaidman 1988).

Outside of the special case $|\Psi'\rangle = |\Psi\rangle$, the DvN axioms reviewed in Subsection 2.2 do not assign a weak value A_w any obvious meaning or interpretation. A glance at the definition (26) does not inspire confidence. A generic weak value is typically complex-valued, and is not a single-system observable, such as the numerical outcome in the spectrum of any observable.[8] Also, a weak value depends on a *pair* of distinct state vectors that are not related to each other by the sort of unitary time evolution characteristic of a closed system. Moreover, a weak value only exists contingently,

this paper's notion of an ensemble observable, probabilities are ensemble observables.

[7] Indeed, some references write weak values *in general* using the bracket notation $\langle A\rangle_w$ (for example, Wiseman 2007; Aharonov, Popescu, Rohrlich, Skrzypczyk 2013; Cohen, Pollak 2018).

[8] In a 2014 reply to the critical analysis of Ferrie and Combes (2014), Vaidman wrote: "[A w]eak value of a variable A is a property of a single quantum system pre-selected in a state $|\psi\rangle$ and post-selected in a state $|\phi\rangle$" (Vaidman 2014), but post-selection is manifestly an ensemble concept and not a single-system concept.

because it is undefined if the two state vectors in its definition have vanishing inner product. Nor is a generic weak value a probability or an expectation value, both of which are ensemble observables that *refer* to single-system observables.

From its definition (26), a generic weak value is quite clearly an ensemble observable that does not refer to a single-system observable, at least as far as the DvN axioms are concerned. There might well be some new axiomatic foundation for quantum theory on which a generic weak value could be a single-system observable, or at least an ensemble observable that *refers* to a single-system observable, but any such interpretation of a weak value would appear to require constructing a new such axiomatic foundation.

Standard experimental methodologies, to be reviewed in Subsection 3.3, extract weak values in only an approximate way, and depend crucially on the suspect use of post-selection, which is arguably not an innocent or innocuous practice (Barandes 2026). For present purposes, it will be useful to present a less practical but exact operational procedure for obtaining a weak value that does not involve post-selection.

Given a quantum system, let A be a self-adjoint operator representing one of the system's observables, let $|\Psi\rangle$ be any unit-norm vector in the system's Hilbert space, and define a corresponding self-adjoint, idempotent projection operator

$$P_\Psi \equiv |\Psi\rangle\langle\Psi| = P_\Psi^\dagger = P_\Psi^2, \qquad \mathrm{tr}\, P_\Psi = 1. \tag{29}$$

Letting z be an arbitrary complex number with complex conjugate z^*, and letting $\mathbb{1}$ denote the identity operator on the system's Hilbert space, define a new self-adjoint, positive-semidefinite operator

$$E(z) \equiv (\mathbb{1} + zA)P_\Psi(\mathbb{1} + z^*A) = E^\dagger(z) \geq 0, \tag{30}$$

which will be called a weak operator.

Suppose now that the system's quantum state is pure and is represented by a state vector $|\Psi'\rangle$ that has nonvanishing inner product with $|\Psi\rangle$, so that $\langle\Psi'|\Psi\rangle \neq 0$. Then from the general formula (24) for the expectation of an observable, the weak operator $E(z)$ has expectation value

$$\langle E(z)\rangle = \langle\Psi'|E(z)|\Psi'\rangle = \langle\Psi'|(\mathbb{1} + zA)|\Psi\rangle\langle\Psi|(\mathbb{1} + z^*A)|\Psi'\rangle. \tag{31}$$

Simplifying, one finds

$$\left.\begin{aligned} \langle E(z)\rangle &= \langle\Psi'|\Psi\rangle\langle\Psi|\Psi'\rangle + z\langle\Psi'|A|\Psi\rangle\langle\Psi|\Psi'\rangle + z^*\langle\Psi|A|\Psi'\rangle\langle\Psi'|\Psi\rangle + |z|^2|\langle\Psi'|A|\Psi\rangle|^2 \\ &= |\langle\Psi'|\Psi\rangle|^2\left[1 + z\frac{\langle\Psi'|A|\Psi\rangle}{\langle\Psi'|\Psi\rangle} + z^*\frac{\langle\Psi|A|\Psi'\rangle}{\langle\Psi|\Psi'\rangle} + |z|^2\left|\frac{\langle\Psi'|A|\Psi\rangle}{\langle\Psi'|\Psi\rangle}\right|^2\right]. \end{aligned}\right\} \tag{32}$$

In terms of the weak value $A_w = A_w(\Psi', \Psi)$, as defined in (26), one can write the expectation value of the weak operator $E(z)$ more succinctly as

$$\langle E(z)\rangle = |\langle\Psi'|\Psi\rangle|^2\big[1 + zA_w + z^*A_w^* + |z|^2|A_w|^2\big]. \tag{33}$$

Variously choosing the complex number z to be 1, -1, i, and $-i$ in the formula (33), one finds

$$\left.\begin{aligned} \langle E(1)\rangle &= |\langle\Psi'|\Psi\rangle|^2\big[1+2\mathrm{Re}\,A_w+|A_w|^2\big],\\ \langle E(-1)\rangle &= |\langle\Psi'|\Psi\rangle|^2\big[1-2\mathrm{Re}\,A_w+|A_w|^2\big],\\ \langle E(i)\rangle &= |\langle\Psi'|\Psi\rangle|^2\big[1-2\mathrm{Im}\,A_w+|A_w|^2\big],\\ \langle E(-i)\rangle &= |\langle\Psi'|\Psi\rangle|^2\big[1+2\mathrm{Im}\,A_w+|A_w|^2\big]. \end{aligned}\right\} \tag{34}$$

It is then a straightforward exercise to solve for $\mathrm{Re}\,A_w$ and $\mathrm{Im}\,A_w$:

$$\mathrm{Re}\,A_w = \mathrm{Re}\,A_w\big(\Psi',\Psi\big) = \frac{\langle E(1)-E(-1)\rangle}{4|\langle\Psi'|\Psi\rangle|^2} = \frac{\langle(1/2)(AP_\Psi+P_\Psi A)\rangle}{|\langle\Psi'|\Psi\rangle|^2}, \tag{35}$$

$$\mathrm{Im}\,A_w = \mathrm{Im}\,A_w\big(\Psi',\Psi\big) = \frac{\langle E(-i)-E(i)\rangle}{4|\langle\Psi'|\Psi\rangle|^2} = \frac{\langle(1/2i)(AP_\Psi-P_\Psi A)\rangle}{|\langle\Psi'|\Psi\rangle|^2}. \tag{36}$$

These two formulas, at least in principle, provide an operational way to extract a weak value experimentally, without any approximations, in terms of the empirical quantities $\langle(1/2)(AP_\Psi+P_\Psi A)\rangle$, $\langle(1/2i)(AP_\Psi-P_\Psi A)\rangle$, and $|\langle\Psi'|\Psi\rangle|^2$. That point is worth emphasizing: weak values can be obtained, at least indirectly, through measurements. In particular, the formulas (35) and (36) involve the expectation values of observables that are represented by the curious self-adjoint combinations $(1/2)(AP_\Psi+P_\Psi A)$ and $(1/2i)(AP_\Psi-P_\Psi A)$, which coincide with the "flux" and "commutator" operators introduced by Cohen, Pollack (2018, equations 3 and 4). However, neither of the two formulas (35) and (36) *as a whole* is the expectation value of an observable, due to the presence of the state-dependent quantity $|\langle\Psi'|\Psi\rangle|^2=\langle\Psi'|P_\Psi|\Psi'\rangle$ in each denominator.

Again, one therefore sees manifestly that a weak value, whether in its entirety as a complex number or in terms of its real and imaginary parts, does not correspond to a single-system observable, nor to the kind of emergent observable that *refers* to a single-system observable, at least on the DvN axioms for quantum theory. At this point in the discussion, the DvN axioms do not assign a weak value *any* obvious or clear physical meaning at all.

Given the preceding analysis, the null hypothesis should be that weak values do not have any physical or metaphysical meaning, and are merely sometimes-convenient complex-valued ratios that are amenable to extraction via appropriate experimental protocols. To claim nonetheless that weak values have some specific interpretation, let alone an interpretation that makes sense at the level of a single system, would be an extraordinary assertion, and the burden should be on the claimants to provide a rigorous argument in favor of that view. Appeals that run afoul of the fallacies reviewed in Subsection 2.1 are insufficient for that purpose, as are appeals to the size of the research literature on weak values.

It is worth being clear at this point on what these sorts appeals often look like, and why they are fallacious. As explained above, it is true that one can extract weak values indirectly through measurements on ensembles. However, the mere fact that one can extract weak values experimentally in this way implies nothing about whether they are single-system observables or even whether they *refer* to single-system observables. In particular, to identify weak values as single-system observables anyway would be to commit the ensemble fallacy (1), to infer anything about a quantum system or an ensemble of quantum systems from weak values obtained via post-selection would be to commit the post-selection fallacy (2), and to conclude anything about the physical interpretation of weak values based solely on the fact that they can be extracted experimentally ("If we can measure it experimentally, then how could we be wrong about its meaning?") would be to commit the measurementist fallacy (3).

3.2 Weak Values for Projectors

Subsection 3.1 has presented reasons for skepticism about assigning a physical significance to weak values. One can strengthen the case for skepticism by focusing attention on a specific class of observables: projection operators, which arguably have the clearest meaning among all the observables of a quantum system, at least to the extent that *any* observables of a quantum system have a clear meaning.

Consider a quantum system with an N-dimensional Hilbert space, and let $|c_i\rangle$ for $i = 1, \ldots, N$ make up the orthonormal eigenbasis for an observable C, so that

$$C|c_i\rangle = c_i|c_i\rangle, \qquad \langle c_i|c_j\rangle = \delta_{ij}. \tag{37}$$

One can then define a corresponding projection-valued measure, or PVM, by

$$P_i \equiv |c_i\rangle\langle c_i| = P_i^\dagger = P_i^2. \tag{38}$$

As a PVM, these projection operators satisfy the conditions of mutual exclusivity (14),

$$P_i P_j = \delta_{ij} P_i \tag{39}$$

and completeness (16),

$$\sum_{i=1}^{N} P_i = \sum_{i=1}^{N} |c_i\rangle\langle c_i| = \mathbb{1}, \tag{40}$$

where $\mathbb{1}$ is the identity operator on the system's Hilbert space. One can then write the observable C as the spectral decomposition

$$C = \sum_{i=1}^{N} c_i P_i = \sum_{i=1}^{N} c_i |c_i\rangle\langle c_i|. \tag{41}$$

Given a state vector $|\Psi\rangle$ for the system, the expectation value (22) of a PVM element P_i has the interpretation of being the probability $p(c_i)$ with which a measurement of the observable C will yield the specific eigenvalue c_i:

$$\langle P_i\rangle = \langle\Psi|P_i|\Psi\rangle = p(c_i). \tag{42}$$

Accordingly, this expectation value $\langle P_i\rangle$ is always real-valued, always lies between the extreme values

0 and 1, and satisfies the usual normalization constraint, in virtue of the completeness relation (40):

$$\sum_{i=1}^{N} p(c_i) = \sum_{i=1}^{N} \langle P_i \rangle = \langle \Psi | \sum_{i=1}^{N} P_i | \Psi \rangle = \langle \Psi | \mathbb{1} | \Psi \rangle = 1. \tag{43}$$

One then interprets the equation $\langle P_i \rangle = 1$ as assigning the value 'true' to the proposition "A measurement of C will yield the eigenvalue c_i," and the equation $\langle P_i \rangle = 0$ as assigning the value 'false' to that same proposition.

These interpretative assertions are consistent with the expectation value $\langle C \rangle$ of C, given the spectral decomposition (41). Indeed,

$$\langle \Psi | C | \Psi \rangle = \sum_{i=1}^{N} c_i \langle \Psi | P_i | \Psi \rangle = \sum_{i=1}^{N} c_i p(c_i) = \langle C \rangle, \tag{44}$$

in line with the definition (22) of an expectation value.

These interpretative assertions and relationships break down for weak values. It is certainly true that the weak values $P_{i,w}$ sum to 1, a feature that bears a resemblance to the normalization constraint (43):

$$\sum_{i=1}^{N} P_{i,w} = \frac{\langle \Psi' | \sum_i P_i | \Psi \rangle}{\langle \Psi' | \Psi \rangle} = \frac{\langle \Psi' | \mathbb{1} | \Psi \rangle}{\langle \Psi' | \Psi \rangle} = 1. \tag{45}$$

However, each individual weak value $P_{i,w}$ can be essentially any complex number, with unbounded modulus, so that 0 and 1 are no longer extreme values. As acknowledged by Vaidman (1996, p. 903, equation 10), the weak value of a projection operator can even be -1. Any connection between projection operators and probabilities (even conditional probabilities) is then utterly lost, as is any reason to say that $P_{i,w} = 1$ should refer to anything being 'true' or that $P_{i,w} = 0$ should refer to anything being 'false.'

As the present work will review, especially in Subsection 3.5, a common view in the research literature is to assert that a weak value $A_w = A_w(\Psi', \Psi)$ should somehow be regarded as the value that the observable A has for an ensemble of quantum systems pre-selected to have the initial state vector $|\Psi\rangle$ and post-selected to have a post-measurement state vector $|\Psi'\rangle$. (See, for example, Aharonov, Vaidman 1990.) What could it possibly mean for a quantum system to have a value for a projection operator that is given by a complex number lying far outside the unit circle from the origin of the complex plane?

Looking back at the weak operator defined in (30) offers little help. For $A = P_i$, one has weak operators defined by

$$E_i(z) \equiv (\mathbb{1} + zP_i) P_\Psi (\mathbb{1} + z^* P_i). \tag{46}$$

The self-adjoint combinations appearing in (35) and (36) then become $(1/2)(P_i P_\Psi + P_\Psi P_i)$ and $(1/2i)(P_i P_\Psi - P_\Psi P_i)$. At least on the DvN axioms, there is no sense in which the first combination having the expectation value $1 \times |\langle \Psi' | \Psi \rangle|^2$ and the second having the expectation value 0, so that $P_{i,w} = \mathrm{Re}\, P_{i,w} + i \mathrm{Im}\, P_{i,w} = 1$, should imply anything about whether any property of the quantum system is 'true.' Nor is there any sense in which both combinations having expectation value 0, so that $P_{i,w} = \mathrm{Re}\, P_{i,w} + i\mathrm{Im}\, P_{i,w} = 0$, should imply that anything about the quantum system is 'false.'

These observations directly challenge statements made, for instance, by Aharonov, Popescu, Rohrlich, Skrzypczyk (2013, pp. 5–6, equations 6—9). The present work will address some of the downstream implications of these arguments in Subsection 4.1.

3.3 The AAV Experimental Protocol

The AAV paper (Aharonov, Albert, Vaidman 1988) presented a practical, if approximate, means of extracting weak values by a somewhat complicated procedure that involves a subtle interplay of entanglement, weak measurements, and ensembles obtained through post-selection. The AAV experimental protocol will be presented here, both to connect the present work with the AAV paper more comprehensively, and also to make clear why the AAV experimental protocol does not, in the end, lead to a more perspicuous interpretation of weak values, despite what one might have hoped.

The derivation to follow will be similar to the derivation of the ABL rule presented in other work (Barandes 2026), but will be fully self-contained.

To begin, one considers three quantum systems: a subject system to be studied, a measuring device, and an external agent or observer. The total system's Hilbert space is given by the tensor product

$$\mathcal{H}_{\mathrm{tot}} = \mathcal{H} \otimes \mathcal{H}_{\mathrm{dev}} \otimes \mathcal{H}_{\mathrm{obs}}, \tag{47}$$

where $\mathcal{H}$, $\mathcal{H}_{\mathrm{dev}}$, and $\mathcal{H}_{\mathrm{obs}}$ are the respective Hilbert spaces for the subject system, the measuring device, and the observer.

As with the ABL rule, one supposes that for the initial state vector $|\Psi_{\mathrm{tot}}\rangle$ of the total system, the subject system has a generic initial state vector $|\Psi\rangle$, and the observer is assigned its usual 'ready' state vector $|\mathrm{obs}(\emptyset)\rangle$. The measuring device, by contrast, is assumed to begin with a specially chosen state vector $|\Phi_{\mathrm{dev}}\rangle$, to be defined momentarily. Altogether, the initial state vector for the total system is

$$|\Psi_{\mathrm{tot}}\rangle = |\Psi\rangle \otimes |\Phi_{\mathrm{dev}}\rangle \otimes |\mathrm{obs}(\emptyset)\rangle. \tag{48}$$

The overall plan will be to pick an observable A belonging to the subject system, together with a second state vector $|\Psi'\rangle$ for the subject system to be used for post-selection, and then indirectly obtain the weak value $A_w(\Psi', \Psi)$ corresponding to A, $|\Psi\rangle$, and $|\Psi'\rangle$, as defined in (26)

One assumes here that the measuring device has a canonical pair of observables, q and p, where the momentum observable p is regarded as the measuring device's pointer variable, and q is its conjugate coordinate.[9] For a standard von Neumann measurement (von Neumann 1932), the measurement outcome would ordinarily be read directly off of the pointer variable p, after imposing a collapse (25) to the total system's state vector. That procedure is modified for the experimental

[9] Note that Bohm (1952), Vaidman (1996a), and Dressel et al. (2014) choose a convention in which the measuring device's canonical coordinate q, rather than its canonical momentum p, is taken to be the pointer variable. For that convention, one should replace q with p in the interaction Hamiltonian (52), and the overall minus sign should be removed.

extraction of weak values.

From the assumption that p is the measuring device's pointer variable, one assumes that it has the same measurement units as A, in which case the product qA has units of the (reduced) Planck constant $\hbar$.

One takes the initial state vector $|\Phi_{\text{dev}}\rangle$ for the measuring device to have a narrow Gaussian profile in coordinate space, centered at $q = 0$. That is, the measuring device's coordinate-space wave function is assumed to have the form

$$\langle q|\Phi_{\text{dev}}\rangle \propto e^{-q^2/4\sigma_q^2}, \tag{49}$$

so that the corresponding coordinate-space probability distribution is

$$|\langle q|\Phi_{\text{dev}}\rangle|^2 \propto e^{-q^2/2\sigma_q^2}, \tag{50}$$

where the standard deviation σ_q is assumed to be very small relative to the reciprocal spacing between the eigenvalues of the observable A, up to a factor of $\hbar$. This assumed initial state vector then has a very broad Gaussian profile in momentum space,

$$\langle p|\Phi_{\text{dev}}\rangle \propto e^{-p^2/4\sigma_p^2}, \tag{51}$$

with a standard deviation $\sigma_p = \hbar/2\sigma_q$ that is very large compared with the spacing between the eigenvalues of A.

This broadness of the measuring device's initial quantum state in momentum space is understood to mean that the measuring device will effectively interact only very weakly with the subject system, so the interaction is accordingly called a 'weak measurement' (Aharonov, Albert, Casher, Vaidman 1987). The analysis ahead will show explicitly that a weak measurement perturbs the quantum state of the measuring device only in a slight, predictable way, and leaves the quantum state of the subject system essentially unchanged, at least for most practical purposes.

Ignoring any other time evolution in the system, under the assumption that the interaction between the subject system and the measuring device occurs over a very short time interval, and following von Neumann (1932), one takes the interaction Hamiltonian to be[10]

$$H(t) = -g(t)qA. \tag{52}$$

Here $g(t)$ is a function that plays the role of a coupling parameter, has measurement units of inverse-time, has compact support that is sharply peaked at the interaction time, and is normalized to unity:

$$\int dt\, g(t) = 1. \tag{53}$$

[10]In the original German: "...dann bleibt von H nur der für die Messung entscheidende Wechselwirkungs-Energieanteil übrig. Für diesen wählen wir die besondere Form $\frac{h}{2\pi i} q \frac{\partial}{\partial r}$" (von Neumann 1932, Section VI.3, p. 236).

Choosing an arbitrary orthonormal basis for the subject system labeled by b, the total system then has the following time evolution:

$$\left.\begin{aligned}|\Psi_{\text{tot}}\rangle &= |\Psi\rangle \otimes |\Phi_{\text{dev}}\rangle \otimes |\text{obs}(\emptyset)\rangle \\ &\mapsto e^{-(i/\hbar)\int dt\, H(t)}|\Psi\rangle \otimes |\Phi_{\text{dev}}\rangle \otimes |\text{obs}(\emptyset)\rangle \\ &= \sum_b \int dq\, [|b\rangle \otimes |q\rangle][\langle b| \otimes \langle q|]e^{-(i/\hbar)\int dt\, H(t)}[|\Psi\rangle \otimes |\Phi_{\text{dev}}\rangle] \otimes |\text{obs}(\emptyset)\rangle.\end{aligned}\right\} \tag{54}$$

Thus, one finds an inner product that needs to be calculated, involving just the subject system and the measuring device:

$$[\langle b| \otimes \langle q|]e^{-(i/\hbar)\int dt\, H(t)}[|\Psi\rangle \otimes |\Phi_{\text{dev}}\rangle] = \langle b|e^{iqA/\hbar}|\Psi\rangle\langle q|\Phi_{\text{dev}}\rangle. \tag{55}$$

Because the measuring device is assumed to have a coordinate-space profile $\langle q|\Phi_{\text{dev}}\rangle \propto \exp\left(-q^2/4\sigma_q^2\right)$ in (55) that is sharply peaked at $q = 0$, one can effectively treat q as a small parameter and therefore approximate the exponential operator as the first-order term in its Taylor series:

$$e^{iqA/\hbar} \approx 1 + \frac{iqA}{\hbar}. \tag{56}$$

Hence,

$$\left.\begin{aligned}\langle b|e^{iqA/\hbar}|\Psi\rangle &\approx \langle b|\left(1 + \frac{iqA}{\hbar}\right)|\Psi\rangle \\ &= \langle b|\Psi\rangle + \frac{iq}{\hbar}\langle b|A|\Psi\rangle \\ &= \langle b|\Psi\rangle\left(1 + \frac{iqA_w}{\hbar}\right) \\ &\approx \langle b|\Psi\rangle e^{iqA_w/\hbar},\end{aligned}\right\} \tag{57}$$

where the weak value A_w is defined as in (26) by

$$A_w = A_w(b, \Psi) \equiv \frac{\langle b|A|\Psi\rangle}{\langle b|\Psi\rangle} \qquad \text{[compare to AAV's eq. (6)]}, \tag{58}$$

and, again, is generically complex-valued. Hence, the inner product (55) reduces approximately to

$$\langle b|\Psi\rangle e^{iqA_w/\hbar}\langle q|\Phi_{\text{dev}}\rangle \approx \langle b|\Psi\rangle\langle q|\Phi_{\text{dev}|A_w}\rangle, \tag{59}$$

where the updated state vector $|\Phi_{\text{dev}|A_w}\rangle$ of the measuring device has coordinate-space wave function

$$\left.\begin{aligned}\langle q|\Phi_{\text{dev}|A_w}\rangle &= e^{iqA_w/\hbar}\langle q|\Phi_{\text{dev}}\rangle \\ &\propto e^{-q^2/4\sigma_q^2 + iqA_w/\hbar}.\end{aligned}\right\} \tag{60}$$

Thus, the total system's state vector (54) becomes approximately

$$\sum_b \int dq \, \langle b|\Psi\rangle \langle q|\Phi_{\mathrm{dev}|A_w(b,\Psi)}\rangle |b\rangle \otimes |q\rangle \otimes |\mathrm{obs}(\emptyset)\rangle, \tag{61}$$

where the explicit notation $A_w(b, \Psi)$ is intended to highlight the entanglement between the subject system and the measuring device.

It is important to note that the orthonormal basis labeled by b here is totally arbitrary at this point. No physical feature of the subsystems or the interaction between the measuring device and the subject system has singled out b or depends on b. The measuring device may have been perturbed by the interaction, but not yet in a way that actually depends on the specific weak values $A_w(b, \Psi)$, for the specific basis labeled by b.

The observer then carries out a standard projective measurement on the *subject system* in the basis labeled by b, thereby yielding the following final state vector for the total system, which now features entanglement between all three subsystems:

$$|\Psi'_{\mathrm{tot}}\rangle = \sum_b \int dq \, \langle b|\Psi\rangle \langle q|\Phi_{\mathrm{dev}|A_w(b,\Psi)}\rangle |b\rangle \otimes |q\rangle \otimes |\mathrm{obs}(b)\rangle. \tag{62}$$

Here obs(b) indicates that the observer has obtained the measurement outcome b.

Crucially, it is only at the conclusion of *this step* that the measuring device's quantum state definitively gains a dependence on the weak values $A_w(b, \Psi)$, for the specific orthonormal basis labeled by b. That dependence on $A_w(b, \Psi)$ did not arise merely from the interaction between the measuring device and the subject system, but required the projective measurement by the observer in the basis labeled by b.

Next, from among the possible final measurement results labeled by b, the observer post-selects a *specific* choice $b = \Psi'$. This measurement and post-selection induces a DvN collapse of the total system's state vector, thereby breaking the entanglement and yielding[11]

$$|\Psi'_{\mathrm{tot}}\rangle = |\Psi'\rangle \otimes \left[\int dq \, \langle q|\Phi_{\mathrm{dev}|A_w(\Psi',\Psi)}\rangle |q\rangle\right] \otimes |\mathrm{obs}\big(\Psi'\big)\rangle. \tag{63}$$

Due to the observer's projective measurement and post-selection, the updated quantum state for the measuring device alone now definitively depends on the specific weak value $A_w = A_w(\Psi', \Psi)$,

[11]As Aharonov and Vaidman note in a separate paper (Aharonov, Vaidman 1990), the observer could instead measure the measuring device's pointer variable p or its conjugate coordinate q first, and then, secondarily, measure the subject system's label b and post-select on $b = \Psi'$, without any meaningful difference to the final state vector or statistics. Whatever the order of these two final steps, it is only due to the observer's choice of b and post-selection on $b = \Psi'$ that any dependence arises on a specific weak value $A_w(b, \Psi) = A_w(\Psi', \Psi)$.

and is given by

$$\left.\begin{aligned} |\Phi_{\mathrm{dev}|A_w}\rangle &= \int dq\,|q\rangle\langle q|\Phi_{\mathrm{dev}|A_w}\rangle \\ &\propto \int dq\,|q\rangle e^{-q^2/4\sigma_q^2+iqA_w/\hbar} \\ &\propto \int dp\,|p\rangle e^{-(p-A_w)^2/4\sigma_p^2}. \end{aligned}\right\} \tag{64}$$

Decomposing the weak value $A_w = \mathrm{Re}\,A_w + i\mathrm{Im}\,A_w$ into its real and imaginary parts, and carrying out a straightforward calculation, one can show that the probability distribution in momentum space turns out to be

$$|\langle p|\Phi_{\mathrm{dev}|A_w}\rangle|^2 \propto e^{-(p-\mathrm{Re}\,A_w)^2/2\sigma_p^2}, \tag{65}$$

whereas the probability distribution in coordinate space is

$$|\langle q|\Phi_{\mathrm{dev}|A_w}\rangle|^2 \propto e^{-\left(q+2\sigma_q^2\mathrm{Im}\,A_w/\hbar\right)^2/2\sigma_q^2}. \tag{66}$$

By assumption, the standard deviation σ_p of the momentum-space probability distribution (65) for the measuring device is very large, so the peak around the expectation value $\langle p\rangle = \mathrm{Re}\,A_w$ is extremely broad, thereby necessitating a large ensemble to gain experimental access to $\mathrm{Re}\,A_w$. Although the standard deviation $\sigma_q = \hbar/2\sigma_p$ of the coordinate-space probability distribution (66) is, by contrast, very small, meaning that the peak around the expectation value $\langle q\rangle = -2\sigma_q^2\mathrm{Im}\,A_w/\hbar = -\hbar\,\mathrm{Im}\,A_w/2\sigma_p^2$ is very narrow, this expectation value itself is extremely small, and so, again, would require a large ensemble to determine experimentally.

These conclusions hold even if, say, the observable A in question is a projector $P = P^\dagger = P^2$ representing a true-or-false question, and even if, furthermore, its weak value happens to be $P_w = 1$ or $P_w = 0$. As explained in Subsection 3.2, there is, in particular, no sense in which the ratio $P_w = \langle\Psi'|P|\Psi\rangle/\langle\Psi'|\Psi\rangle$ having the value 1 means that anything about the subject system is definitively 'true,' or having the value 0 means that anything about the subject system is definitively 'false,' again in contrast with claims made by Aharonov, Popescu, Rohrlich, Skrzypczyk (2013, pp. 5–6, equations 6–9), as noted in Subsection 3.2.

In principle, then, to obtain a weak value like A_w, one sets up a large physical ensemble of $N \gg 1$ identically prepared bipartite systems. Each bipartite system consists of a subject system with an observable A and initially represented by a state vector $|\Psi\rangle$, together with a measuring device prepared in a narrow Gaussian in the space of the coordinate conjugate to its pointer variable. One then lets the two subsystems in each bipartite system suitably interact. At the end of the whole experiment, an external agent or observer carries out a projective measurement on each subject system and keeps only the bipartite systems for which the subject system is found to have some specified final state vector $|\Psi'\rangle$, thereby deliberately implementing a 'post-selection' that projects the subject system's quantum state down to $|\Psi'\rangle$. The observer then carries out a projective measurement on each measuring device itself—specifically, either a measurement of the measuring device's pointer variable p or its conjugate coordinate q. Having carried out this overall experimental protocol on a large physical ensemble $N \gg 1$ of these bipartite systems, and after computing the relevant statistical means, the external agent can obtain an estimate of the real and imaginary parts of the weak value A_w.

Despite predating the AAV paper, the AACV paper considered the slightly more general possibility of implementing the extraction of more than one weak value in the course of the overall experimental protocol. If each measuring device has multiple pointer variables $p_1, \ldots, p_n$, with corresponding conjugate coordinates $q_1, \ldots, q_n$, then one can extract weak values in succession on the same subject system, all from the same pre-selected state vector $|\Psi\rangle$ and the same post-selected state vector $|\Psi'\rangle$. For this purpose, one replaces the original Hamiltonian (52) with

$$\sum_{\alpha=1}^{n} H_\alpha(t) = -\sum_{\alpha=1}^{n} g_\alpha(t) q_\alpha A_\alpha, \tag{67}$$

and one assumes that the measuring device's coordinate-space profile is a Gaussian that is sharply peaked at $(q_1, \ldots, q_n) = (0, \ldots, 0)$:

$$\langle q_1, \ldots, q_n | \Phi_{\mathrm{dev}} \rangle \propto e^{-\sum_\alpha q_\alpha^2 / 4\sigma_{q_\alpha}^2}. \tag{68}$$

In place of the inner product (55), one now has

$$[\langle b| \otimes \langle q_1, \ldots, q_n|] e^{-(i/\hbar)\sum_\alpha \int dt\, H_\alpha(t)} [|\Psi\rangle \otimes |\Phi_{\mathrm{dev}}\rangle] = \langle b| e^{i\sum_\alpha q_\alpha A_\alpha/\hbar} |\Psi\rangle \langle q_1, \ldots, q_n | \Phi_{\mathrm{dev}} \rangle. \tag{69}$$

Treating the coordinates $q_1, \ldots, q_n$ as effectively small parameters—a crucial step for treating the interactions as weak measurements—one then obtains

$$\langle b| e^{i\sum_\alpha q_\alpha A_\alpha/\hbar} |\Psi\rangle \approx \langle b| \left(1 + \frac{i}{\hbar} \sum_{\alpha=1}^{n} q_\alpha A_\alpha \right) |\Psi\rangle \approx \langle b|\Psi\rangle e^{i\sum_\alpha q_\alpha A_{\alpha,w}/\hbar}, \tag{70}$$

so

$$\langle b| e^{i\sum_\alpha q_\alpha A_\alpha/\hbar} |\Psi\rangle \langle q_1, \ldots, q_n | \Phi_{\mathrm{dev}} \rangle \approx \langle b|\Psi\rangle \langle q_1, \ldots, q_n | \Phi_{\mathrm{dev}|A_{1,w},\ldots,A_{n,w}} \rangle. \tag{71}$$

Here the measuring device's updated coordinate-space wave function generalizes (60) to

$$\langle q_1, \ldots, q_n | \Phi_{\mathrm{dev}|A_{1,w},\ldots,A_{n,w}} \rangle \propto e^{-\sum_\alpha \left(q_\alpha^2/4\sigma_{q_\alpha}^2 + i q_\alpha A_{\alpha,w}/\hbar\right)}. \tag{72}$$

After post-selection to $b = \Psi'$, the weak values are, in keeping with (26), given by

$$A_{\alpha,w} \equiv \frac{\langle \Psi'|A_\alpha|\Psi\rangle}{\langle \Psi'|\Psi\rangle} = A_{\alpha,w}\big(\Psi', \Psi\big) \quad [\alpha = 1, \ldots, n]. \tag{73}$$

These weak values are all independent of each other, and so are not subject to conditions like the uncertainty principle. They also satisfy a form of robustness or replicability, in the sense that if the same observable is subjected to a weak-value measurement twice, then the corresponding weak values will be the same:

$$A_\alpha = A_\beta \implies A_{\alpha,w} = A_{\beta,w}. \tag{74}$$

3.4 The Interpretation of Weak Values

A weak value A_w, as defined in (26), manifestly involves a renormalized matrix element $\langle\Psi'|A|\Psi\rangle/\langle\Psi'|\Psi\rangle$ of the self-adjoint operator A. Being able to calculate matrix elements of self-adjoint operators through an ensemble-based experimental protocol is clearly a useful technique on its own terms, again as reviewed by Dressel et al. (2014), but one might ask for more. In particular, given the discussion in Subsection 3.1 and the review of the AAV experimental protocol in Subsection 3.3, one might reasonably wonder if weak values have a more profound physical or metaphysical significance.

However, as explained in Subsection 3.1, the DvN axioms identify a weak value only as an irreducibly emergent observable of an *ensemble* of systems, and not even the kind of ensemble observable that *refers* to a single-system observable. Nothing about the AAV experimental protocol changes that basic fact.

It is true that in the derivation presented in Subsection 3.3, the final state vector $|\Phi_{\text{dev}|A_w}\rangle$ of each measuring device, as expressed in (64), contains a dependence on the weak value $A_w = A_w(\Psi', \Psi)$. However, merely recording an ensemble observable's value into the internal memory of a single measuring device does not make that ensemble observable into a single-system observable, or else *all* ensemble observables would be single-system observables. What matters is not the system into which the observable's value is *stored*, but the system from which the observable is *obtained*. Generic weak values, just like DvN measurement probabilities and expectation values, are *obtained* from ensembles of subject systems, and not from any one subject system.

It might be tempting to argue that weak values must have some sort of single-system meaning based on the fact that the AAV experimental protocol involves measurement-like interactions between the subject systems and the measuring devices, together with the fact that these measurement-like interactions perturb the quantum states of the measuring devices (Vaidman 1996a). If a procedure perturbs the quantum states of measuring devices, then how can whatever produces the perturbation fail to exist in some physical sense? The question, however, is about whether or not whatever produces the perturbation is a single-system feature or property, and, moreover, whether or not it is a statistical artifact of the choice of post-selection.

One should keep in mind that the quantum states of the measuring devices did *not* develop their dependence on the weak value $A_w = A_w(\Psi', \Psi)$ due only to the interactions between the measuring devices and the subject systems. As emphasized explicitly in Subsection 3.3 immediately before and after (62), that dependence on $A_w = A_w(\Psi', \Psi)$ was the result of the external observer carrying out a final projective measurement and post-selection on the subject systems for the specific choice of state vector $|\Psi'\rangle$. Those decisions and actions taken by the observer were what imbued the quantum states of the measuring devices with a dependence on $|\Psi'\rangle$, and, consequently, with a dependence on the weak value $A_w = A_w(\Psi', \Psi)$.

To argue that the weak value A_w was just there, waiting to be found, would be akin to arguing that a sculpture already exists in a block of marble and merely needs to be revealed, as imagined by Robert Allyn in this paper's epigraph.[12]

In particular, attempting to identify a weak value obtained from the AAV experimental protocol as directly implying some value for a single-system observable, or neglecting the manifest dependence of a weak value on the arbitrary choice of post-selection by the observer, would be to engage in the sort of fallacious reasoning discussed in Subsection 2.1.

An example might be helpful here. Forgetting about quantum theory for a moment, consider a large ensemble of classical pendulums, each one adjacent to its own individual detector and electronic memory-storage device. Suppose that each detector measures the amplitude, frequency, and phase of its nearby pendulum, so that the detector's memory-storage device develops a classical correlation with that nearby pendulum. Suppose also that an external experimenter has the ability to observe the pendulums directly, bypassing the detectors.

By a judicious choice of post-selection directly on the pendulums, the experimenter can cull the overall ensemble down to a subensemble of pendulums whose amplitudes, frequencies, and phases compose the Fourier series for Bach's Toccata and Fugue. As a side-effect, the memory-storage devices in this subensemble—which, again, are each correlated with their nearby pendulum—would then collectively contain the Toccata and Fugue, as a Fourier series. This experimental protocol would therefore provide an indirect way to encode the Toccata and Fugue into memory-storage devices for various practical uses.

It would obviously be incorrect, however, to take this result to imply anything about the individual properties of the pendulums, let alone to ascribe a small amount of 'Toccata-and-Fugue-ness' ('fugacity'?) to each pendulum. Bach's Toccata and Fugue here is clearly an emergent ensemble observable, and a statistical artifact of the experimenter's choice of post-selection on the pendulums, rather than a reflection of a property intrinsic to each individual pendulum itself. To confuse this ensemble observable with a single-system property of each pendulum, or to use this ensemble observable to make direct inferences about any single-system property of each pendulum, would entail committing the ensemble fallacy (1). Moreover, to draw an erroneous conclusion about the pendulums based on the use of post-selection would mean committing the post-selection fallacy (2). Finally, to attempt to evade these two charges of fallacious reasoning merely by pointing out that the Toccata and Fugue can, in fact, be extracted experimentally from this protocol ("How could the result be mistaken if we can get it from an experiment?") would be to commit the measurementist fallacy (3).

Granted, the foregoing example is, by assumption, classical. However, if one wanted to construct a defense of conventional interpretations of weak values against the charges of fallacious reasoning presented in this paper, it would not be enough merely to point out that weak values arise from quantum-mechanical experiments, and to say that quantum systems are different from classical systems. Of course it is true that quantum systems are different from classical systems. Quantum systems are, after all, characterized by the DvN axioms (or axioms that at least make empirically similar claims), with exotic-looking downstream implications that appear to be conceptually distinct

[12] Variants of this quotation are sometimes attributed to Michelangelo.

from the sorts of phenomena that one finds for classical systems. However, a proper defense of conventional interpretations of weak values would require carefully showing that quantum theory differs from classical physics in the *right kinds of ways* that make *the right kind of difference.* Indeed, putting aside the classical example presented above, the criticisms laid out in the present work and the fallacies listed in Subsection (2.1) are not based in classical physics, but take weak values seriously from a quantum-mechanical perspective.

3.5 Historical Approaches to the Interpretation of Weak Values

In a 2007 paper, Wiseman referred to a weak value as a kind of 'mean' value, purportedly akin to the expectation value of a single-system observable:

> Consider a weak measurement of some observable a. A *weak value*, denoted $\langle \hat{a}_w \rangle_{|\psi\rangle}$, is the *mean* value from such weak measurements on an ensemble of systems, each prepared in the state $|\psi\rangle$. So far, the weak value is no different from the strong value, the ensemble mean value of strong or precise measurements. However, the weak value differs from the strong value if one calculates the mean from a subensemble obtained by post-selecting only those results for which a later strong measurement reveals the system to be in state $|\phi\rangle$. It is convenient to denote such a weak value by ${}_{\langle\phi|}\langle \hat{a}_w \rangle_{|\psi\rangle}$." [Wiseman 2007, p. 4, emphasis in the original]

As the quoted passage indicates, Wiseman even went so far as to use bracket notation in general for weak values, as a way of mimicking the usual notation for an expectation value. This bracket notation for weak values was also used by Aharonov, Popescu, Rohrlich, Skrzypczyk (2013), and by Cohen, Pollak (2018). The identification of weak values with means or averages occurs in other papers as well, including in a review article by Dressel et al. (2014), which calls them "conditioned averages."

However, the expectation value $\langle A \rangle$ of a single-system observable A does not get its interpretation as a mean value for a single-system observable from the AAV experimental procedure of extracting weak values, as laid out in Subsection 3.3. Instead, $\langle A \rangle$ gets its fundamental interpretation as a mean value over the possible values a in the spectrum $\sigma(A)$ of the single-system observable A from the basic definition (22),

$$\langle A \rangle \equiv \sum_a a\, p(a), \tag{75}$$

which is manifestly a probability-weighted average of the possible values a of the single-system observable A, where $p(a)$ is the Born-rule probability for the measurement outcome a.

It is true that if an expectation value like $\langle A \rangle$ is, practically speaking, *extracted* via the AAV experimental protocol, then that protocol will involve averaging over the distribution (65) of the measuring device's pointer variable and averaging over the distribution (66) of the measuring device's conjugate coordinate. Nevertheless, this merely *practical technique* of averaging over an ensemble of measuring-device variables is conceptually different from the *fundamental definition* of $\langle A \rangle$ as an average or mean over the *single-system observable* A. Moreover, for a weak value $A_w(\Psi', \Psi)$ for which $|\Psi'\rangle \not\propto |\Psi\rangle$, this average or mean over an ensemble of *measuring–device variables* is the *only* sense in which the weak value is a mean value, at least according to the DvN axioms. These subtly distinct senses of the word 'mean' are yet another reason why there have been so many debates over the proper interpretation of weak values.

To make clear why these conceptually different notions of 'mean' really matter, and are not merely a form of hairsplitting, it might help to consider a more familiar-looking example. Suppose that a teacher running a classroom of some fixed number of students decides to assign each student a distinct natural number, chosen at random. The teacher then writes these numbers down on a sheet of paper, but does not show the list to the students. Later on, the teacher could certainly 'observe' or 'measure' this list of arbitrary natural numbers and compute their statistical mean, but the answer would be a mean only in the narrow sense of being an average over measurement results, without being an average over single-student attributes possessed by the students themselves. As such, this statistical mean would reveal nothing whatsoever about the students. By contrast, if the teacher instead decides to measure each student's height, or decides to ask the students for their ages, then the teacher could compute statistical means of a qualitatively different conceptual character, because a mean of heights or a mean of ages would not merely be an average of measurement results, but would be an average over single-student attributes. In particular, a classroom visitor looking at the statistical mean of the list of arbitrary natural numbers would learn nothing about the students, but the visitor could learn something interesting from the mean of heights or from the mean of ages.

The AACV paper focused on instrumentalist questions of measurement, and said very little about the interpretation of weak values beyond that. The AAV paper likewise described weak values mostly as the possible outcomes of a particular class of experimental protocols.

In a 1990 paper by Aharonov and Vaidman titled "Properties of a Quantum System During the Time Interval Between Two Measurements" (Aharonov, Vaidman 1990), one began to see stronger claims of interpretation. The paper's title itself suggested that a generic weak value should be understood as a property of a single quantum system. Indeed, the paper's third paragraph included the following statements:

> The most important outcome of our approach is the possibility to define a new concept: the *weak value* of a quantum variable. It is a physical property of a quantum system between two measurements, i.e., a property of a system belonging to an ensemble that is both preselected and postselected. This property can manifest itself through a measurement that fulfills certain requirements of weakness. [Ibid., p. 11, emphasis in the original]

Note the explicit inference of a single-system property merely from the fact that the system belongs to a specially chosen ensemble defined through post-selection. Again, the attempt to identify a single-system property or observable with an ensemble observable, or to infer the former from the latter, is precisely an example of the ensemble fallacy (1). The attempt to make direct inferences about systems through the deliberate use of post-selection implies committing the post-selection fallacy (2).

The paper tried to get around these obstructions by considering the extraction of weak values in a *single* measurement, but on a large composite system consisting of $N \gg 1$ individual spin-1/2 subsystems—which is, of course, an ensemble from the perspective of each single spin-1/2 system. Even then, the paper concluded that the measurement outcome could have been described as the result of measuring a traditional observable of the large composite system according to the standard DvN formalism: “We shall now show how the above result can be explained using the standard formalism, which we in no way dispute.” (Aharonov, Vaidman 1990, p. 16)

In a 1991 paper, Aharonov and Vaidman showed that for certain choices of pre-selected state vector $|\Psi\rangle$, observable A, and post-selected state vector $|\Psi'\rangle$, the weak value A_w coincided with one of the observable's eigenvalues (Aharonov, Vaidman 1991). This result did not, however, establish that weak values were single-system observables. Indeed, given certain choices of quantum state, an expectation value $\langle A\rangle$ can sometimes be equal to one of the eigenvalues of the corresponding observable A, but that does not make expectation values single-system observables. A quantity that is only *contingently* related to a single-system observable, in a state-dependent way, is still categorically different from a single-system observable.

In a 1996 paper, Vaidman, inspired by the definition of an “element of reality” from the famous EPR paper (Einstein, Podolsky, Rosen 1935), introduced a new definition, intended to capture weak values:

> I suggest to take this property to be the definition: *If we are certain that a procedure for measuring a certain variable will lead to a definite shift of the unchanged probability distribution of the pointer, then there is an element of reality: the variable equal to this shift.* [Vaidman 1996a, p. 898, emphasis in the original]

It is certainly true that the full AAV experimental protocol for extracting a weak value leads to shifts in quantum states of measuring devices, as (64) makes clear. Moreover, the question over whether there are elements of reality distinct from quantum states is an important topic in quantum foundations and in debates over the interpretation of quantum theory. However, this definition of an element of reality does not distinguish between single-system elements of reality and ensemble elements of reality, nor does it address the way that post-selection by an external observer is necessary to produce the desired weak value, because, again, as emphasized in the discussion surrounding (62), the interactions between the measuring devices and the subject systems alone are insufficient. Those issues are key to the critique of the research literature on weak values laid out in the present work. Indeed, as Vaidman added:

> In such a case, a measurement performed on a single system does not yield the value of the shift (the element of reality), but such measurements performed on a large enough ensemble of identical systems yield the shift with any desirable precision. [Ibid., p. 898]

In the conclusion of the paper, Vaidman went on to acknowledge that weak values were not generically single-system observables, but nonetheless made the case for taking them seriously:

> I certainly see a deficiency of weak-measurement elements of reality defined above in the situations in which they cannot be measured on a single system. Still, I do not think that the fact that a weak value cannot be measured on a single system prevents it from being a "reality." We know that the measuring device shifts its pointer exactly according to the weak value, even though we cannot find it because of the large uncertainty of the pointer position. We can verify this knowledge performing measurements on an ensemble. [Ibid., p. 904]

Weak values have been measured in many experiments in the years since the AACV and AAV paper were published. In that regard, the AACV and AAV papers represented a marvelous triumph. Nevertheless, without committing the measurementist fallacy (3), there is no reason to think that the success of those experiments in obtaining matrix elements (26) provides any support for the interpretation of weak values presented in the subsequent research literature as single-system observables, or as providing a way to infer single-system observables.

4 Questionable Applications of Weak Values

4.1 Quantum Cheshire Cats

In 2013, Aharonov, Popescu, Rohrlich, and Skrzypczyk (APRS) published a paper titled "Quantum Cheshire Cats" in the *New Journal of Physics* (Aharonov, Popescu, Rohrlich, Skrzypczyk 2013). The APRS paper claimed that the location of a *single photon* could seemingly be separated from the photon's polarization. Taking a cue from Lewis Carroll's *Alice in Wonderland*, whose Cheshire Cat could be separated from its grin, the paper called the photon a "quantum Cheshire Cat."

The grounds for this claim were counterfactual reasoning with post-selection, together with the invocation of weak values. Although the APRS paper acknowledged that these notions all applied at the level of an ensemble of many particles, either measured individually or all at once, the paper nonetheless committed the ensemble fallacy (1) by explicitly claiming to make single-particle inferences from these ensemble observables. Indeed, in the paper's introduction, one found these statements:

> Yet, as we will show here, in the curious way of quantum mechanics, photon polarization may exist where there is no photon at all. At least this is the story that quantum mechanics tells via measurements on a pre- and post-selected ensemble. [Ibid., p. 2]

The APRS paper went on to add:

> Let us first ask which way the photon went inside the interferometer. We will show that, given the pre- and post-selection, *with certainty the photon went through the left arm.* [...] Thus the non-demolition measurement in the right arm never finds the photon there, indicating that the photon must have gone through the left arm. [...] *The Cat is therefore in the left arm. But can we find its grin elsewhere?* [...] We will discover that

> *there is angular momentum in the right arm.* [...] We seem to see what Alice saw—a grin without a cat! We know with certainty that the photon went through the left arm, yet we find angular momentum in the right arm. [Ibid., pp. 3–4, emphasis in the original]

The use of language like "the photon went through the left arm" and "The Cat is therefore in the left arm" attributes *happening* meanings to properties of quantum systems that are distinct from measurement statements, and therefore runs directly into the category problem described in Subsection 2.2. Moreover, the statements in this quoted passage are all single-particle statements, and are therefore inadmissible due to committing the ensemble fallacy (1).

The APRS paper did not address these issues, but instead suggested a different problem:

> But could this conclusion really be right? It is, ultimately, open to the following criticism. We never actually simultaneously measured the location and the angular momentum. Indeed, our conclusions above were reached by measuring location on some photons and angular momentum on others. The immediate implication is that all we have here is a paradox of counterfactual reasoning [...]. That is, we have made statements about where the photon is, and about where the angular momentum is, that are paradoxical as long as we don't actually perform all the relevant measurements simultaneously. [Ibid., p. 4, emphasis in the original]

The APRS paper claimed to solve this problem by appealing instead to weak values of projectors having the special values 1 or 0, but, as explained in the present work, weak values do not provide an escape from the ensemble fallacy (1) or from the post-selection fallacy (2). Furthermore, as explained in Subsection 3.2, the values 1 and 0 do not have the intuitive meanings of 'true' and 'false' for the weak values of projectors.

Other papers since 2013 have commented on quantum Cheshire Cats both theoretically and experimentally (for example, Denkmayr et al. 2014, Corrêa et al. 2015, Duprey et al. 2018), but none have approached the issue according to the specific form of critical analysis carried out in the present work.

4.2 Bohmian Trajectories

Weak values have also shown up in the research literature on Bohmian mechanics (de Broglie 1930; Bohm 1952a, 1952b), a hidden-variables formulation or interpretation of quantum theory in which the Hilbert-space ingredients of textbook quantum theory are augmented with classical-like particles that have definite positions at all times and follow hidden trajectories that are guided or piloted by their overall wave function.

In 2007, Wiseman published a paper titled "Grounding Bohmian Mechanics in Weak Values and Bayesianism" (Wiseman 2007) in which he argued that weak values could be used to assign an objective physical reality to those hidden trajectories by giving an operational definition to a specific choice of vector field:

> Since velocity is defined by the rate of change of position, in fact we want simply to make a weak measurement of initial position, then a strong measurement of position a short time τ later. We then have the following *operational definition* for the velocity for a particle at position $\mathbf{x}$:
>
> $$\mathbf{v}(\mathbf{x};t) \equiv \lim_{\tau \to 0} \tau^{-1} E[\mathbf{x}_{\text{strong}}(t+\tau) - \mathbf{x}_{\text{weak}}(t) | \mathbf{x}_{\text{strong}}(t+\tau) = \mathbf{x}]. \qquad \text{[Wiseman's eq. (5)]} \tag{76}$$
>
> Here $E[a|F]$ denotes the average of a over the (post-selected) ensemble where F is true. [Ibid., p. 4, emphasis in the original]

It is true that the ensemble-level vector field $\mathbf{v}(\mathbf{x};t)$, as defined above, has measurement units of a velocity, but Wiseman explicitly attributes it to "a particle at position $\mathbf{x}$," thereby triggering the ensemble fallacy (1). Moreover, the weak value $\mathbf{x}_{\text{weak}}(t)$ implicitly depends on the post-selection condition $\mathbf{x}_{\text{strong}}(t+\tau) = \mathbf{x}$, so to make direct inferences about the particle based on this post-selection condition would mean committing the post-selection fallacy (2). Finally, the weak value $\mathbf{x}_{\text{weak}}(t)$ itself does not have a clear physical interpretation, as the present work has argued, so the difference between $\mathbf{x}_{\text{strong}}(t+\tau) - \mathbf{x}_{\text{weak}}(t)$ does not have a clear physical meaning, either.

An interesting feature of the vector field $\mathbf{v}(\mathbf{x};t)$ defined in (76) is that, as Wiseman explains, it coincides with the definition of the velocity of a Bohmian particle located at position $\mathbf{x}$ and time t for a suitable choice of Hamiltonian. However, if this observation is used to justify the definition (76), then (76) cannot then be used to justify the physical reality of Bohmian trajectories, on pain of logical circularity.

To avoid this logical circularity, Wiseman attempts to provide an independent justification for (76):

> Consider a naive experimentalist, with no knowledge of QM beyond the following basic facts about experiments at the microscopic scale: (i) no matter how carefully a preparation procedure is repeated, the measured properties of the particle will vary in different runs of the experiment; (ii) if a given property of the particle is measured strongly (arbitrarily accurately) then in general this will drastically alter the future distribution of measurement results; (iii) if an arbitrarily weak measurement is used instead, the future distribution can remain essentially unaltered. For such an experimentalist, equation (5) would be the only sensible way to measure the velocity of a particle at position $\mathbf{x}$. Thus I will call the velocity in equation (5) the *naively observable velocity*, and contend that it is the most natural operational definition of velocity. [Ibid., p. 5, emphasis in the original]

The first two desiderata (i) and (ii) are certainly facts about presently-known kinds of experiments at the microscopic scale, and are fully in keeping with the DvN axioms, as laid out in Subsection 2.2.

The third desideratum (iii), however, is not a fact, because a fact cannot rely fundamentally on an ill-defined concept. Notice that (iii) depends on the notion of a "weak measurement," as described in Subsection 3.3, together with the implicit assumption that the weak value extracted from a weak measurement has some sort of clear physical meaning. Without being able to attach a clear physical meaning to a weak value, especially in light of the dangers of post-selection that are involved in constructing one, (iii) does not offer any means of obtaining any other physical notions, let alone the notion of a single-particle velocity. To argue otherwise by appealing to the fact that the vector field $\mathbf{v}(\mathbf{x};t)$ consists of quantities that can be obtained experimentally, and that indeed *have* been obtained in later experimental work (Kocsis et al. 2011; Mahler et al. 2016), would be to commit the measurementist fallacy (3). (For further arguments against the interpretation of $\mathbf{v}(\mathbf{x};t)$ as describing single-particle velocities, see Fankhauser, Dürr 2021.)

5 Conclusion

The present work provided a critical analysis of interpretational claims made about weak values, arguing that while weak values are clearly useful for practical applications, like signal amplification and state tomography, weak values do not reveal interesting physical or metaphysical properties of individual quantum systems. Since the first appearance of weak values in the research literature, there were great hopes that they could provide a new window into the inner workings of quantum theory. Unfortunately, nature has not been as forthcoming with its secrets as one might have wished.

As in other work (Barandes 2026), a larger issue is the use of post-selection in quantum-foundations papers. As statisticians know, post-selection can lead to spurious correlations and other statistical artifacts. Distinguishing those sorts of statistical artifacts from truly quantum behavior can be challenging. As a general rule, then, papers that deliberately deploy post-selection in the course of deriving exotic results should be especially careful to explain why those results are really due to quantum theory itself. Research journals should be sure to hold authors to that standard.

Acknowledgments

The author would like to thank David Albert, David Kagan, Logan McCarty, Xiao-Li Meng, Filip Niewinski, John Norton, James Robins, Yossi Sirote, Kelly Werker Smith, Lev Vaidman, Larry Wasserman, Michael Weissman, and Ken Wharton for helpful discussions.

References

[1] Y. Aharonov, D. Z. Albert, A. Casher, and L. Vaidman. “Surprising Quantum Effects”. *Physics Letters A*, 124(4-5):199–203, September 1987. `doi:10.1016/0375-9601(87)90619-0`.

[2] Y. Aharonov, D. Z. Albert, and L. Vaidman. “How the Result of a Measurement of a Component of the Spin of a Spin-1/2 Particle Can Turn Out to be 100”. *Physical Review Letters*, 60(14):1351, April 1988. `doi:10.1103/PhysRevLett.60.1351`.

[3] Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz. "Time Symmetry in the Quantum Process of Measurement". *Physical Review*, 134(6B):B1410, June 1964. doi:10.1103/PhysRev.134.B1410.

[4] R. Allyn. "Whittier's Poems". *The Methodist Quarterly Review*, pages 72–92, January 1858. Book review of *The Poems of John Greenleaf Whittier*.

[5] Y. Aharonov, S. Popescu, D. Rohrlich, and P. Skrzypczyk. "Quantum Cheshire Cats". *New Journal of Physics*, 15(11):113015, November 2013. arXiv:1202.0631, doi:10.1088/1367-2630/15/11/113015.

[6] Y. Aharonov and D. Rohrlich. "Comment on 'How the Result of a Single Coin Toss Can Turn Out to be 100 Heads'", 2014. URL: https://arxiv.org/abs/1410.0381, arXiv:1410.0381, doi:10.48550/arXiv.1410.0381.

[7] Y. Aharonov and L. Vaidman. "Properties of a Quantum System During the Time Interval Between Two Measurements". *Physical Review A*, 41(1):11, January 1990. doi:10.1103/PhysRevA.41.11.

[8] Y. Aharonov and L. Vaidman. "Complete Description of a Quantum System at a Given Time". *Journal of Physics A: Mathematical and General*, 24(10):2315, January 1991. doi:10.1088/0305-4470/24/10/018.

[9] G. Bacciagaluppi. "The Role of Decoherence in Quantum Mechanics". In E. N. Zalta and U. Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2025 edition, 2025. URL: https://plato.stanford.edu/archives/spr2025/entries/qm-decoherence.

[10] J. A. Barandes. "The Stochastic-Quantum Correspondence". *Philosophy of Physics*, 3(1):8, June 2025. arXiv:2302.10778, doi:10.31389/pop.186.

[11] J. A. Barandes. "The ABL Rule and the Perils of Post-Selection". 2026. URL: https://arxiv.org/abs/2602.07402, arXiv:2602.07402, doi:10.48550/arXiv.2602.07402.

[12] J. S. Bell. "On the Problem of Hidden Variables in Quantum Mechanics". *Reviews of Modern Physics*, 38(3):447–452, July 1966. doi:10.1103/RevModPhys.38.447.

[13] J. Berkson. "Limitations of the Application of Fourfold Table Analysis to Hospital Data". *Biometrics Bulletin*, 2(3):47–53, June 1946. URL: http://www.jstor.org/stable/3002000.

[14] D. J. Bohm. "A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables. I". *Physical Review*, 85(2):166–179, January 1952. doi:10.1103/PhysRev.85.166.

[15] D. J. Bohm. "A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables. II". *Physical Review*, 85(2):180–193, January 1952. doi:10.1103/PhysRev.85.180.

[16] A. Brodutch. "Comment on "How the Result of a Single Coin Toss Can Turn Out to be 100 Heads"". *Physical Review Letters*, 114(11):118901, March 2015. arXiv:1410.8510, doi:10.1103/PhysRevLett.114.118901.

[17] R. Corrêa, M. França Santos, C. H. Monken, and P. L. Saldanha. "'Quantum Cheshire Cat' as Simple Quantum Interference". *New Journal of Physics*, 17(5):053042, May 2015. arXiv:1409.0808, doi:10.1088/1367-2630/17/5/053042.

[18] E. Cohen. "A Comment on "How the Result of a Single Coin Toss Can Turn Out to be 100 Heads"", 2014. URL: https://arxiv.org/abs/1409.8555, arXiv:1409.8555, doi:10.48550/arXiv.1409.8555.

[19] E. Cohen and E. Pollak. "Determination of Weak Values of Quantum Operators Using Only Strong Measurements". *Physical Review A*, 98(4):042112, October 2018. arXiv:1804.11298, doi:10.1103/PhysRevA.98.042112.

[20] S. R. Cole, R. W. Platt, E. F. Schisterman, H. Chu, D. Westreich, D. Richardson, and C. Poole. "Illustrating Bias Due to Conditioning on a Collider". *International Journal of Epidemiology*, 39(2):417–420, November 2009. arXiv:https://academic.oup.com/ije/article-pdf/39/2/417/18480441/dyp334.pdf, doi:10.1093/ije/dyp334.

[21] L. de Broglie. *An Introduction to the Study of Wave Mechanics*. E. P. Dutton and Company, Inc., 1930.

[22] T. Denkmayr, H. Geppert, S. Sponar, H. Lemmel, A. Matzkin, J. Tollaksen, and Y. Hasegawa. "Observation of a Quantum Cheshire Cat in a Matter-Wave Interferometer Experiment". *Nature Communications*, 5:4492, July 2014. arXiv:1312.3775, doi:10.1038/ncomms5492.

[23] P. A. M. Dirac. *The Principles of Quantum Mechanics*. Oxford University Press, 1st edition, 1930.

[24] Q. Duprey, S. Kanjilal, U. Sinha, D. Home, and A. Matzkin. "The Quantum Cheshire Cat Effect: Theoretical Basis and Observational Implications". *Annals of Physics*, 391:1–15, April 2018. arXiv:1703.02959, doi:10.1016/j.aop.2018.01.011.

[25] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd. "Colloquium: Understanding Quantum Weak Values: Basics and Applications". *Reviews of Modern Physics*, 86(1):307–316, March 2014. arXiv:1305.7154, doi:10.1103/RevModPhys.86.307.

[26] A. Einstein, B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?". *Physical Review*, 47(10):777–780, May 1935. doi:10.1103/PhysRev.47.777.

[27] C. Ferrie and J. Combes. "How the Result of a Single Coin Toss Can Turn Out to be 100 Heads". *Physical Review Letters*, 113(12):120404, September 2014. arXiv:1403.2362, doi:10.1103/PhysRevLett.113.120404.

[28] C. Ferrie and J. Combes. "Ferrie and Combes Reply:". *Physical Review Letters*, 114(11):118902, March 2015. doi:10.1103/PhysRevLett.114.118902.

[29] J. Fankhauser and P. M. Dürr. "How (Not) to Understand Weak Measurements of Velocities". *Studies in History and Philosophy of Science Part A*, 85:16–29, February 2021. arXiv:2309.10395, doi:10.1016/j.shpsa.2020.12.002.

[30] A. Hájek. "Interpretations of Probability". In E. N. Zalta and U. Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2023 edition, 2023. URL: https://plato.stanford.edu/archives/win2023/entries/probability-interpret.

[31] M. A. Hernán, S. Hernández-Díaz, and J. M. Robins. "A Structural Approach to Selection Bias". *Epidemiology*, 15(5):615–625, September 2004. doi:10.1097/01.ede.0000135174.63482.43.

[32] J. R. Hance, J. Rarity, and J. Ladyman. "Weak Values and the Past of a Quantum Particle". *Physical Review Research*, 5(2):023048, April 2023. arXiv:2109.14060, doi:10.1103/PhysRevResearch.5.023048.

[33] S. Kocsis, B. Braverman, S. Ravets, M. J. Stevens, R. P. Mirin, L. K. Shalm, and A. M. Steinberg. "Observing the Average Trajectories of Single Photons in a Two-Slit Interferometer". *Science*, 332(6034):1170–1173, June 2011. doi:10.1126/science.1202218.

[34] S. B. Kochen and E. P. Specker. "The Problem of Hidden Variables in Quantum Mechanics". *Indiana University Mathematics Journal*, 17(1):59–87, 1967.

[35] D. H. Mahler, L. Rozema, K. Fisher, L. Vermeyden, K. J. Resch, H. M. Wiseman, and A. Steinberg. "Experimental Nonlocal and Surreal Bohmian Trajectories". *Science Advances*, 2(2):e1501466–e1501466, February 2016. doi:10.1126/sciadv.1501466.

[36] A. Romito, A. N. Jordan, Y. Aharonov, and Y. Gefen. "Weak Values are Quantum: You Can Bet on It". *Quantum Studies: Mathematics and Foundations*, 3(1):1–4, 2016. arXiv:1508.06304, doi:10.1007/s40509-015-0069-z.

[37] R. Shankar. *Principles of Quantum Mechanics*. Plenum Press, 2nd edition, 1994.

[38] L. Vaidman. "Weak-Measurement Elements of Reality". *Foundations of Physics*, 26(7):895–906, July 1996. arXiv:quant-ph/9601005, doi:10.1007/BF02148832.

[39] L. Vaidman. "Comment on "How the Result of a Single Coin Toss Can Turn Out to be 100 Heads"", 2014. URL: https://arxiv.org/abs/1409.5386, arXiv:1409.5386, doi:10.48550/arXiv.1409.5386.

[40] J. von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Berlin: Springer, 1932.

[41] H. M. Wiseman. "Grounding Bohmian Mechanics in Weak Values and Bayesianism". *New Journal of Physics*, 9(6):165, June 2007. arXiv:0706.2522v1, doi:10.1088/1367-2630/9/6/165.