Foundations of Physics (2018) 48:1568–1589
https://doi.org/10.1007/s10701-018-0216-6

Quantum Theory and the Limits of Objectivity

Richard Healey[1]

Received: 27 June 2018 / Accepted: 11 September 2018 / Published online: 19 September 2018

Abstract
Three recent arguments seek to show that the universal applicability of unitary quantum theory is inconsistent with the assumption that a well-conducted measurement always has a definite physical outcome. In this paper I restate and analyze these arguments. The import of the first two is diminished by their dependence on assumptions about the outcomes of counterfactual measurements. But the third argument establishes its intended conclusion. Even if every well-conducted quantum measurement we ever make will have a definite physical outcome, this argument should make us reconsider the objectivity of that outcome.

Keywords Quantum theory · Objectivity · Wigner's friend · Brukner · Frauchiger and Renner

1 Introduction

Quantum theory is taken to be fundamental to contemporary physics in large part because countless measurements have yielded outcomes that conform to its predictions. Experimenters take great care to ensure that each quantum measurement has an outcome that is not just a subjective impression but an objective, physical event. However, in the continuing controversy in quantum foundations QBists [1,2] and others [3–5] have come to question and even deny the principle that a well-conducted quantum measurement has a definite, objective, physical outcome. This principle should not be

✉ Richard Healey
rhealey@email.arizona.edu

1 University of Arizona, Tucson, AZ, USA

abandoned lightly: objective data provide the platform on which scientific knowledge rests.[1] We should demand a water-tight argument before giving it up.

In this paper I analyze three recent arguments that quantum theory, consistently applied, entails that not every quantum measurement can have a definite, objective, physical outcome. I say 'can have', not 'has', because each argument requires a *Gedankenexperiment* far more extreme even than that of Schrödinger's cat. The first two arguments' dependence on questionable implicit assumptions severely limits their significance. But I think the third argument at least succeeds in deflating a certain ideal of objectivity in the quantum domain. I assume throughout that an outcome of a quantum measurement is definite only if it is unique—an assumption rejected by Everettians such as Deutsch [7] and Wallace [8]. Assuming the objectivity of a physical outcome, an Everettian may take an argument like these considered here as offered in support of that outcome's non-uniqueness, as suggested by the title of [9].

2 Brukner's Argument

Brukner's argument [3,4] applies Bell's theorem [10] to an extension of Wigner's [11] friend scenario. My restatement of the most recent version [4] of his argument renames Brukner's characters and introduces clarifying notation.

Before describing his own *Gedankenexperiment*, Brukner considers Deutsch's [7] twist on Wigner's original friend scenario. So consider first a scenario in which Zeus[2] is contemplating possible measurements on Xena's otherwise physically isolated lab X, inside which Xena has measured the z-spin of a single spin-1/2 particle 1 prepared in the superposed state in the z-spin basis

$$|x\rangle_1 = 1/\sqrt{2}(|\uparrow\rangle_1 + |\downarrow\rangle_1) \tag{1}$$

Assuming the universal applicability of unitary quantum mechanics, Zeus assigns to the combined system $1X$ after Xena's measurement the entangled state

$$|\Phi\rangle_{1X} = 1/\sqrt{2}(|\uparrow\rangle_1 \, |\text{“ up”}\rangle_X + |\downarrow\rangle_1 \, |\text{“down”}\rangle_X), \tag{2}$$

where $|\text{“up”}\rangle_X$ (for example) represents a state in which if Zeus were to observe the contents of Xena's lab he would certainly (with probability 1) find her reporting the outcome of her measurement of z-spin on particle 1 as $+\hbar/2$ and that her recording device had indeed recorded that value.

[1] Even if no item of data is so certain as to be immune from rejection in the light of further scientific investigation. Recall Popper's [6, p. 94] famous metaphor:
"Science does not rest upon solid bedrock. The bold structure of its theories rises, as it were, above a swamp. It is like a building erected on piles. The piles are driven down from above into the swamp, but not down to any natural or 'given' base; and if we stop driving the piles deeper, it is not because we have reached firm ground. We simply stop when we are satisfied that the piles are firm enough to carry the structure, at least for the time being."

[2] Brukner calls this character Wigner, but I have reserved that name for another character with analogous powers.

Zeus can try to verify his assignment of state $|\Phi\rangle_{1X}$ by performing a measurement of a dynamical variable A_x represented by the operator $\hat{A}_x$ on $H_1 \otimes H_X$

$$\hat{A}_x = |\uparrow\rangle_1 |\text{“ up”}\rangle_X \langle\downarrow|_1 \langle\text{“ down”}|_X + |\downarrow\rangle_1 |\text{“ down”}\rangle_X \langle\uparrow|_1 \langle\text{“ up”}|_X . \quad (3)$$

This measurement will (with probability 1) yield the outcome $+1$ while leaving the state $|\Phi\rangle_{1X}$ undisturbed. After this successful verification, Zeus's apparatus and memory establish the truth of statement A_x^+: “Zeus's outcome is $A_x = +1$” and the falsity of A_x^-: “Zeus's outcome is $A_x = -1$.” But despite his entangled state assignment, Zeus may have some reason to believe that Xena has indeed observed a definite outcome of her measurement of z-spin.

As Deutsch [7] pointed out, no violation of unitary quantum theory is involved if Xena passes a message out of her lab to Zeus reporting that she has seen a definite outcome, as long as this contains no information about what that outcome was. Zeus may try to see for himself whether Xena has seen a definite outcome by performing his own measurement on her lab and its contents, of a dynamical variable A_z represented by the operator $\hat{A}_z$ on $H_1 \otimes H_X$

$$\hat{A}_z = |\uparrow\rangle_1 |\text{“ up”}\rangle_X \langle\uparrow|_1 \langle\text{“ up”}|_X - |\downarrow\rangle_1 |\text{“ down”}\rangle_X \langle\downarrow|_1 \langle\text{“ down”}|_X . \quad (4)$$

If Zeus's outcome is $A_z = +1$ he may judge this to verify the statement A_z^+: “Xena's outcome is z^+,” and falsify A_z^-: “Xena's outcome is z^-”, while outcome $A_z = -1$ reverses these judgments. These judgments are not warranted by the (false) assumption that an ideal quantum measurement just faithfully reveals the pre-existing value of the measured variable. Instead, their warrant rests on the assumption that Xena's outcome is accessible to other observers by consulting her records. Failure of such intersubjectivity would undermine Xena's outcome's claim to objectivity, at least in this epistemic sense.

Since the measurement of A_x leaves the state $|\Phi\rangle_{1X}$ unchanged, Zeus may first perform that measurement to establish the truth of A_x^+, then measure A_z to verify the truth of A_z^+ (or, alternatively, of A_z^-). So in this preliminary scenario Zeus has some reason to believe that not only his own measurements but also Xena's measurement had a definite, physical outcome. Moreover, if he measures only A_x he can then pass a message with its outcome to Xena, also without disturbing the state $|\Phi\rangle_{1X}$: so Xena, too, will have reason to believe that both A_x^+ and A_z^+ (or A_z^-) are true and that Zeus's measurement of A_x as well as her own measurement of 1's z-spin had a definite, physical outcome.

Now consider the statements $c(A_x^+)$: “A_x^+ would be true if Zeus were to measure A_x”, $c(A_x^-)$: “A_x^- would be true if Zeus were to measure A_x”; $c(A_z^+)$: “Zeus's outcome would be $A_z = +1$ if he were to measure A_z”, $c(A_z^-)$: “Zeus's outcome would be $A_z = -1$ if he were to measure A_z”. Zeus has reason to believe $c(A_x^+)$ is true and $c(A_x^-)$ is false whether or not he measures A_x, since $|\Phi\rangle_{1X}$ predicts the truth of A_x^+ (with probability 1). Whether or not he measures A_z, Zeus has reason to believe that one of $c(A_z^+), c(A_z^-)$ is true while the other is false in state $|\Phi\rangle_{1X}$. Assuming measurements have definite, objective outcomes, he should take his conditional outcome simply to reflect Xena's actual outcome: Xena got z^+ if and only if Zeus would get $+1$, while

Xena got z^- if and only if Zeus would get − 1. Provided that Xena's measurement had a definite actual outcome it follows that exactly one of $c(A_z^+)$ or $c(A_z^-)$ is true.

After analyzing this preliminary scenario, Brukner [4] introduces his own, more complex, *Gedankenexperiment*. Each of Xena and Yvonne is located in a separate laboratory. These laboratories are initially completely physically isolated, and this isolation is preserved except for the processes specified below. An entangled pair of spin-1/2 particles is prepared, with particle 1 in Xena's lab X and particle 2 in Yvonne's lab Y. In [4] the initial state assigned to 12 (in the z-spin basis) is

$$\begin{aligned} |\psi\rangle_{12} &= -\sin\theta/2\,\left|\phi^+\right\rangle_{12} + \cos\theta/2\,\left|\psi^-\right\rangle_{12}, \text{ where} \\ \left|\phi^+\right\rangle_{12} &= 1/\sqrt{2}(|\uparrow\rangle_1\,|\uparrow\rangle_2 + |\downarrow\rangle_1\,|\downarrow\rangle_2) \\ \left|\psi^-\right\rangle_{12} &= 1/\sqrt{2}(|\uparrow\rangle_1\,|\downarrow\rangle_2 - |\downarrow\rangle_1\,|\uparrow\rangle_2). \end{aligned} \tag{5}$$

Then Xena measures the z-spin of particle 1 in her lab, while Yvonne measures the z-spin of particle 2 in her lab. Assume that the measurement in each laboratory has a definite, physical outcome, registered by a particle detector, recorded in a computer (or on paper) and experienced by Xena or Yvonne respectively.

Each of Zeus and Wigner is also located in a separate laboratory. Xena's laboratory is located wholly within Zeus's, while Yvonne's is located wholly within Wigner's. But to this point each laboratory has remained completely physically isolated insofar as there has been no direct physical interaction between any of these four laboratories.

Assuming (no-collapse) quantum theory is universally applicable, there is a correct quantum state for Zeus and Wigner to assign to the joint physical system consisting of the entire contents of both Xena's and Yvonne's laboratories and this state evolved unitarily throughout the interactions involved in each of their spin-component measurements. (Note that in assigning this state, Zeus and Wigner are here treating Xena and Yvonne themselves as quantum (sub)systems.) Assuming for simplicity that the spin-component measurements were non-disturbing, we may write this joint state after Xena's and Yvonne's measurements as

$$\begin{aligned} |\Psi\rangle_{12XY} &= -\sin\theta/2\,\left|\Phi^+\right\rangle + \cos\theta/2\,\left|\Psi^-\right\rangle, \text{ where} \\ \left|\Phi^+\right\rangle &= 1/\sqrt{2}\left(\left|A_{up}\right\rangle\left|B_{up}\right\rangle + |A_{down}\rangle\,|B_{down}\rangle\right) \\ \left|\Psi^-\right\rangle &= 1/\sqrt{2}\left(\left|A_{up}\right\rangle|B_{down}\rangle - |A_{down}\rangle\left|B_{up}\right\rangle\right). \end{aligned} \tag{6}$$

Here X represents the entire contents of Xena's lab (including Xena) and Y represents the entire contents of Yvonne's lab (including Yvonne), except the measured particles 1, 2. $\left|A_{up}\right\rangle$, $|A_{down}\rangle$ are eigenstates of $\hat{A}_z$.

We may define an analogous pair of self-adjoint operators on $H_2 \otimes H_Y$ as follows:

$$\begin{aligned} \hat{B}_z &= |\uparrow\rangle_2\,|\text{``up''}\rangle_Y\,\langle\uparrow|_2\,\langle\text{``up''}|_Y - |\downarrow\rangle_2\,|\text{``down''}\rangle_Y\,\langle\downarrow|_2\,\langle\text{``down''}|_Y \\ \hat{B}_x &= |\uparrow\rangle_2\,|\text{``up''}\rangle_Y\,\langle\downarrow|_2\,\langle\text{``down''}|_Y + |\downarrow\rangle_2\,|\text{``down''}\rangle_Y\,\langle\uparrow|_2\,\langle\text{``up''}|_Y \end{aligned}$$

where magnitude B_z on $2Y$ uniquely corresponds to $\hat{B}_z$ and B_x to $\hat{B}_x$. The state (6) predicts that the statistics of the (assumed, definite) outcomes of Zeus's and

Wigner's measurements will violate the associated Clauser–Horne–Shimony–Holt inequality

$$S = \langle A_z B_z \rangle + \langle A_z B_x \rangle + \langle A_x B_z \rangle - \langle A_x B_x \rangle \leq 2 \qquad \text{(CHSH)}$$

in which $\langle A_x B_z \rangle$, for example, is the correlation function of a probability distribution for the outcomes of measurements of magnitudes A_x, B_z. The inequality CHSH is violated, for example, by state $|\Psi\rangle_{12XY}$ for which $S = 2\sqrt{2}$ if $\theta = \pi/4$.

Suppose that the sequence of measurements by Xena, Yvonne, Zeus and Wigner is repeated in many trials, with Zeus's measurement of A_x or A_z and Wigner's measurement of B_x or B_z varied randomly and independently from trial to trial. Violation of (CHSH) by statistics collected in a large number of such trials is perfectly consistent with the assumption

> *Definite Outcomes*: In every such trial each of Xena's, Yvonne's, Zeus's and Wigner's measurements has a definite, physical outcome.

The assumption of *Definite Outcomes* does not even make it unlikely that a large number of Zeus's and Wigner's outcomes in repeated trials will display correlations in violation of (CHSH). Indeed the Born rule *predicts* that the outcomes of Zeus's and Wigner's measurements will violate CHSH: if $\theta = \pi/4$ then $S = 2\sqrt{2}$. Why might one think otherwise?

Brukner [4] takes his argument to disprove the following postulate

> **Postulate** ("Observer-independent facts") *The truth-values of the propositions A_i of all observers form a Boolean algebra $\mathcal{A}$. Moreover, the algebra is equipped with a (countably additive) positive measure $p(A) \geqq 0$ for all statements $A \in \mathcal{A}$, which is the probability for the statements to be true.*

To evaluate the bearing of his argument on the assumption of *Definite Outcomes* one must specify propositions purporting to describe such outcomes. Brukner's discussion of the preliminary scenario suggests these include A_z^+, A_z^-, A_x^+, and A_x^-. The symmetry of the *Gedankenexperiment* further suggests they include propositions B_z^+, B_z^-, B_x^+ and B_x^-, each of which states the outcome of an analogous measurement by Yvonne or by Wigner. Is there any reason to believe that application of Brukner's **Postulate** to the propositions $\mathcal{B} = \{A_z^+, A_z^-, B_z^+, B_z^-, A_x^+, A_x^-, B_x^+, B_x^-\}$ yields the promised no-go theorem?

In no repetition are both A_x and A_z measured—the experimental arrangements are mutually exclusive, as are those for B_x and B_z. If A_x is not measured, then neither A_x^+ nor A_x^- describes an actual outcome: and if B_x is not measured, then neither B_x^+ nor B_x^- describes an actual outcome. So unless A_x, B_x are measured in a repetition, the propositions of all observers that describe the actual definite outcomes assumed by *Definite Outcomes* is not the whole of $\mathcal{B}$ but merely a compatible subset $\mathcal{B}^*$ forming a Boolean algebra which may readily be equipped with a (countably additive) positive measure: just use the Born probabilities from state (6) and extend this to each proposition describing the outcome of an actual measurement by Xena or by Yvonne by equating its outcome to that of the corresponding measurement by Zeus or by Wigner (so, for example, A_z^+ is true if and only if the outcome of Zeus's measurement of A_z is $A_z = +1$, and both propositions have the same probability).

If A_x, B_x are measured in a repetition, the propositions of all observers describing the actual definite outcomes assumed by *Definite Outcomes* is the whole of $\mathcal{B}$. But the propositions $\{A_x^+, A_x^-, B_x^+, B_x^-\}$ form a Boolean algebra whose structure is respected by the obvious truth-assignments, and the Born probabilities from (6) define a probability measure on this algebra. In the absence of any further constraints it is easy to extend this truth-assignment and probability measure to the full algebra $\mathcal{B}$.

So the assumed actual outcomes in each trial can certainly be described by propositions A_i of all observers forming a Boolean algebra $\mathcal{A}$. Moreover, this algebra may be equipped with a (countably additive) positive measure $p(A) \geqq 0$ for all statements $A \in A$, which may be taken as the probability for the statements to be true in that trial. A no-go theorem is not derivable through the application of *Definite Outcomes* to propositions A_i of all observers that describe their actual outcomes in any, or all, repetitions of Brukner's *Gedankenexperiment*.

What happens if instead the "propositions of all observers" concern not their actual but their hypothetical outcomes? Consider the set $c(\mathcal{B}) = \{c(A_z^+), c(A_z^-), c(B_z^+), c(B_z^-), c(A_x^+), c(A_x^-), c(B_x^+), c(B_x^-)\}$ of subjunctive conditionals describing the outcomes of hypothetical measurements. Assume that if such a measurement is actually made in a trial then the corresponding conditional has the same truth-value as its consequent (so, for example, if A_x is measured with outcome $A_x = +1$ then $c(A_x^+)$ is true as well as A_x^+). Unlike the simpler scenario discussed earlier, when A_x^+, A_x^- are replaced by the corresponding subjunctive statements $c(A_x^+)$, $c(A_x^-)$: "If A_x were measured then the definite outcome would be $+1$ (-1)", in state (6) there is no reason to suppose that either of these statements even *has* a truth-value if Zeus does not measure A_x. Nor should $c(B_x^+)$, $c(B_x^-)$ be expected to have truth-values when Wigner does not measure B_x.

Unless A_x,B_x are both measured in a trial, replacement of propositions about actual definite outcomes of a measurement by such conditionals fails to generate a Boolean algebra of propositions of all observers whose truth-value assignment respects that algebra. But quantum theory predicts violation of the inequality CHSH only for the outcomes of actual measurements. Because of the physical incompatibility of Zeus's joint measurement of A_x and A_z and of Wigner's measurement of B_x and B_z, these predictions must concern four distinct kinds of trials, which is what necessitated variation of measurements by Zeus and by Wigner from trial to trial. *Definite Outcomes* implies that the set $c(\mathcal{B})$ forms a Boolean algebra whose structure is respected by a joint truth-assignment and is equipped with a (countably additive) positive measure *at most* in the case of a repetition in which Zeus measures A_x and Wigner measures B_x. So the violation of that inequality in state (6) does not refute *Definite Outcomes. As it stands,* Brukner's argument [3,4] provides no good reason to doubt that every quantum measurement has a definite, objective, physical outcome.

In correspondence, Brukner has proposed a slight modification that avoids this objection and promises to strengthen the argument. In the modified scenario, Zeus measures A_x and Wigner measures B_x in *every* trial. As in the simpler Wigner's friend scenario, Zeus may appeal to the epistemic objectivity of Xena's outcome to infer that

$c(A_z^+)$ has the same truth-value as A_z^+, and $c(A_z^-)$ has the same truth-value as A_z^-.[3] Since *Definite Outcomes* implies that one of A_z^+, A_z^- is true and the other false, it then follows that in every repetition one of $c(A_z^+)$, $c(A_z^-)$ is true and the other false, even though Zeus actually measures A_x and not A_z in that repetition. Similarly, in every repetition one of $c(B_z^+)$, $c(B_z^-)$ is true and the other is false. So in each repetition the set of propositions $\{c(A_z^+), c(A_z^-), c(B_z^+), c(B_z^-)\}$ always forms a Boolean algebra whose truth-value assignment and probability distribution follow from those assigned to the assumed actual outcomes of Xena's and Yvonne's measurements. This will be true in every trial of this modified scenario.

Definite Outcomes now implies that every proposition in the full algebra $c(\mathcal{B})$ has a truth-value and these truth-values respect the algebra's structure. Moreover, any (countably additive) positive measure $p(A) \geqq 0$ for all statements in $c(\mathcal{B})$ must be constrained by a transformed inequality obtained from CHSH by replacing each reference to an actual outcome by a reference to the corresponding hypothetical outcome (though for A_x,B_x the hypothetical outcome *is* the actual outcome). If quantum theory were to predict violation of this transformed inequality then it would imply that *Definite Outcomes* is false.

But quantum theory predicts probabilities only for the outcomes of actual measurements, and neither A_z nor B_z is actually measured in this modified scenario. Only A_x, B_x and the z-spins of **1**,**2** are measured in each trial, and quantum theory makes no predictions of the joint probability distribution for Zeus's and Yvonne's pairs of measurement outcomes, or that for Wigner's and Xena's pairs of measurement outcomes. This is to be expected, since even if *Definite Outcomes* is true, these outcome pairs are not epistemically accessible by any observer (including the four agents named in this scenario), so their statistics are of no scientific interest.

3 Frauchiger and Renner's Argument

Here is a simplified restatement of the argument of Frauchiger and Renner [9,12]. The appendix compares its strategy to that of the arguments on which it is based and supplies a translation to the notation of [12].

Four physical observers are each located in their own separate laboratories. Every laboratory is initially completed physically isolated, and this isolation is preserved except for the processes specified below. In one laboratory observer Xena has prepared a "biased quantum coin" c in state

$$|ready\rangle_c = \frac{1}{\sqrt{3}} |heads\rangle_c + \frac{\sqrt{2}}{\sqrt{3}} |tails\rangle_c . \tag{7}$$

At time $t = 0$ Xena "flips the coin" by implementing a measurement on c of observable f with orthonormal eigenstates $|heads\rangle_c$, $|tails\rangle_c$ by means of a unitary interaction with c.

[3] Though this inference is now questionable, since in this context the antecedent "Zeus measures A_z" of the counterfactuals $c(A_z^+)$, $c(A_z^-)$ is not merely false but incompatible with Zeus's actual measurement of A_x.

$$|ready\rangle_c \, |ready\rangle_{X^-} \Longrightarrow |\psi\rangle^0_{cX^-} = \frac{1}{\sqrt{3}} \, |heads\rangle^0_c \, |heads\rangle_{X^-} + \frac{\sqrt{2}}{\sqrt{3}} \, |tails\rangle^0_c \, |tails\rangle_{X^-} \,. \tag{8}$$

Here and elsewhere I put a numerical superscript n on a state to mark its unitary evolution up to just after time $t = n$. X^- is a system representing the entire contents of Xena's lab (including Xena herself, but neither c nor a qubit system s whose state she is about to prepare), while X is $X^- + c$. $|heads\rangle_{X^-}$, $|tails\rangle_{X^-}$ are orthonormal eigenstates of a binary indicator observable on X whose eigenvalue $x = 1$ represents Xena's outcome "heads" and whose eigenvalue $x = -1$ represents Xena's outcome "tails".

$$\hat{x} \, |heads\rangle_{X^-} = |heads\rangle_{X^-}$$
$$\hat{x} \, |tails\rangle_{X^-} = - \, |tails\rangle_{X^-} \,. \tag{9}$$

Assume Xena's measurement of f on c has a unique, physical outcome: either "heads" or "tails".

At time $t = 1$, if the outcome was "heads", Xena prepares the state of a qubit system s in her lab in state $|\downarrow\rangle_s$: if the outcome was "tails", Xena prepares s in state $|\rightarrow\rangle_s = 1/\sqrt{2}(|\downarrow\rangle_s + |\uparrow\rangle_s)$. Xena can do this by means of a unitary interaction between s and X, yielding the following state

$$|\psi\rangle^1_{cX^-s} = \frac{1}{\sqrt{3}} \, |heads\rangle^1_c \, |heads\rangle^1_{X^-} \, |\downarrow\rangle_s + \frac{\sqrt{2}}{\sqrt{3}} \, |tails\rangle^1_c \, |tails\rangle^1_{X^-} \, |\rightarrow\rangle_s \tag{10}$$

$$= \frac{1}{\sqrt{3}} \Big(|heads\rangle^1_c \, |heads\rangle^1_{X^-} \, |\downarrow\rangle_s + |tails\rangle^1_c \, |tails\rangle^1_{X^-} \, |\downarrow\rangle_s + |tails\rangle^1_c \, |tails\rangle^1_{X^-} \, |\uparrow\rangle_s \Big) \,. \tag{11}$$

Xena then transfers system s out of her lab and into Yvonne's lab, keeping c in her own lab.

Let Y^- be a system consisting of the entire contents of Yvonne's lab (including Yvonne but not the system s transferred to her by Xena), while Y is $Y^- + s$. At time $t = 2$ Yvonne measures observable S_z on s with orthonormal eigenstates $|\downarrow\rangle^2_s$, $|\uparrow\rangle^2_s$ by means of another unitary interaction within her lab, yielding state

$$|\psi\rangle^2_{cX^-sY^-} = \frac{1}{\sqrt{3}} \begin{pmatrix} |heads\rangle^2_c \, |heads\rangle^2_{X^-} \, |\downarrow\rangle^2_s \, |-1/2\rangle_{Y^-} + \\ |tails\rangle^2_c \, |tails\rangle^2_{X^-} \, |\downarrow\rangle^2_s \, |-1/2\rangle_{Y^-} + \\ |tails\rangle^2_c \, |tails\rangle^2_{X^-} \, |\uparrow\rangle^2_s \, |+1/2\rangle_{Y^-} \end{pmatrix} \tag{12}$$

which we can rewrite as

$$|\psi\rangle^2_{XY} = \frac{1}{\sqrt{3}} (|heads\rangle_X \, |-1/2\rangle_Y + |tails\rangle_X \, |-1/2\rangle_Y + |tails\rangle_X \, |+1/2\rangle_Y). \tag{13}$$

Let y be a binary indicator observable on Y whose eigenvalue $y = 1$ represents Yvonne's outcome "+ 1/2" and whose eigenvalue $y = -1$ represents Yvonne's outcome "− 1/2".

$$\hat{y}\,|{+1/2}\rangle_Y = |{+1/2}\rangle_Y$$
$$\hat{y}\,|{-1/2}\rangle_Y = -\,|{-1/2}\rangle_Y\,. \tag{14}$$

Assume Yvonne's measurement of S_z on s has a unique, physical outcome: either "+ 1/2" or "− 1/2".

The state (13) of XY just after $t = 2$ may also be expressed as

$$|\psi\rangle_{XY}^2 = \frac{1}{\sqrt{3}}(\sqrt{2}\,|fail\rangle_X\,|{-1/2}\rangle_Y + |tails\rangle_X\,|{+1/2}\rangle_Y) \tag{13a}$$
$$= \frac{1}{\sqrt{3}}(|heads\rangle_X\,|{-1/2}\rangle_Y + \sqrt{2}\,|tails\rangle_X\,|fail\rangle_Y) \tag{13b}$$
$$= \frac{1}{2\sqrt{3}}(3\,|fail\rangle_X\,|fail\rangle_Y + |fail\rangle_X\,|OK\rangle_Y - |OK\rangle_X\,|fail\rangle_Y + |OK\rangle_X\,|OK\rangle_Y) \tag{13c}$$

where the states $|fail\rangle_X$, $|OK\rangle_X$ are defined by

$$|OK\rangle_X = \frac{1}{\sqrt{2}}(|heads\rangle_X - |tails\rangle_X)$$
$$|fail\rangle_X = \frac{1}{\sqrt{2}}(|heads\rangle_X + |tails\rangle_X) \tag{15}$$

and the states $|fail\rangle_Y$, $|OK\rangle_Y$ are defined by

$$|OK\rangle_Y = \frac{1}{\sqrt{2}}(|{-1/2}\rangle_Y - |{+1/2}\rangle_Y)$$
$$|fail\rangle_Y = \frac{1}{\sqrt{2}}(|{-1/2}\rangle_Y + |{+1/2}\rangle_Y). \tag{16}$$

At time $t = 3$ Zeus measures observable z on X with orthonormal eigenstates $|fail\rangle_X^3$, $|OK\rangle_X^3$ and records a unique, physical outcome: either "fail", or "OK". At time $t = 4$ Wigner measures observable w on Y with orthonormal eigenstates $|fail\rangle_Y^4$, $|OK\rangle_Y^4$ and records a unique, physical outcome: either "fail", or "OK". Finally, at $t = 5$ Wigner consults Zeus and notes the outcome of his measurement of z.

In arriving at the quantum state assignment (13) (and its equivalents), Wigner has correctly applied unitary quantum theory to the specified interactions. Equation (13c) implies that with probability $1/12$ (slightly more than 8%) the outcomes of Zeus's and Wigner's measurements will both be "OK". We now investigate Wigner's reasoning about the outcomes of Xena's and Yvonne's measurements in such a case.

Step 1 At $t = 5$ Zeus tells me that the outcome of his measurement of z on X at $t = 3$ was "OK", so I infer that the unique outcome of his measurement of z on X at $t = 3$ was "OK".

Step 2 Yvonne measured observable S_z on s at time $t = 2$. If her outcome had been "$-1/2$" and not "$+1/2$", then Eq. (13a) implies (with probability 1) that the unique outcome of Zeus's measurement of z on X at $t = 3$ was "fail" and not "OK". But I inferred in step 1 that the unique outcome of his measurement of z on X at $t = 3$ was "OK". So I now infer (with probability 1) that the unique outcome of Yvonne's measurement of observable S_z on s at time $t = 2$ was "$+1/2$".

Step 3 Xena measured observable f on c at $t = 0$. If her outcome had been "heads" and not "tails", then Eq. (13) implies (with probability 1) that the unique outcome of Yvonne's measurement of S_z on s at time $t = 2$ was "$-1/2$" and not "$+1/2$". But I inferred in step 2 that the unique outcome of Yvonne's measurement of observable S_z on s at time $t = 2$ was "$+1/2$". So I now infer (with probability 1) that the unique outcome of Xena's measurement of f on c at $t = 0$ was "tails".

*Step 4** The unique outcome of my measurement of w on Y at $t = 4$ was "OK". But Eq. (13b) implies (with probability 1) that if the unique outcome of Xena's measurement of f on c at $t = 0$ had been "tails", the unique outcome of my measurement of w on Y at $t = 4$ would have been "fail". So I infer that the unique outcome of Xena's measurement of f on c at $t = 0$ was "heads" and not "tails".

Since the conclusion of step 4* contradicts the conclusion of step 3, Wigner's reasoning has here led to a contradiction. The reasoning depended on several assumptions, at least one of which must be rejected to restore consistency. These include the three assumptions:

Universality Quantum theory may be applied to all systems, including macroscopic apparatus, observers and laboratories.

No collapse When an observable is measured on a quantum system in a physically isolated laboratory, the state vector correctly assigned by an external observer to the combined system + laboratory evolves unitarily throughout.

Unique outcome A measurement of an observable has a unique, physical outcome.

Unique outcome corresponds to what Frauchiger and Renner [12] call (S). The appendix discusses the relation between these three assumptions and Frauchiger and Renner's assumptions (C), (Q), and (S). But step 4* depends on an additional assumption that should be questioned and, I argue, rejected:

Intervention insensitivity The truth-value of an outcome-counterfactual is insensitive to the occurrence of a physically isolated intervening event.

An outcome-counterfactual is a statement of the form $O_{t_1} \Box\!\!\rightarrow O_{t_2}$ where O_t states the outcome of a quantum measurement at t, $t_1 < t_2$, and $A \Box\!\!\rightarrow B$ means "If A had been the case then B would have been the case": An event then intervenes just if it occurs in the interval (t_1, t_2), and it is physically isolated if it occurs in a laboratory that is then physically isolated from laboratories where O_{t_1}, O_{t_2} occur.

To see the problem with step 4* of Wigner's reasoning, focus on the outcome-counterfactual "If the unique outcome of Xena's measurement of f on c at $t = 0$ had been "tails", the unique outcome of my measurement of w on Y at $t = 4$ would have been "fail"." Zeus's measurement of z on X at $t = 3$ was an intervening event that occurred in Zeus's laboratory $Z \cup X$ (taken now to encompass the laboratory X on which he performs his measurement of z). At $t = 3$, $Z \cup X$ is still physically isolated from Y and W. So the assumption of *Intervention Insensitivity* would license step 4* of Wigner's reasoning.

But *Intervention Insensitivity* actually conflicts with the other assumptions of the argument. To see why, consider how Wigner should apply quantum theory to Zeus's measurement of z on X at $t = 3$, in accordance with *Universality* and *No collapse*. Equation (13a) implies

$$|\psi\rangle^2_{XY} = \frac{1}{\sqrt{3}}\left[\sqrt{2}\,|fail\rangle_X\,|{-1/2}\rangle_Y + \frac{1}{\sqrt{2}}(|fail\rangle_X - |OK\rangle_X)\,|{+1/2}\rangle_Y\right] \tag{17}$$

Assume for simplicity that Zeus's measurement on X is non-disturbing. Wigner knows that Zeus made a non-disturbing measurement of z on X at $t = 3$. So the state he should assign to XYZ immediately following this measurement is

$$\begin{aligned}
|\psi\rangle^3_{XYZ} &= \frac{1}{\sqrt{3}}\begin{bmatrix} \sqrt{2}\,|fail\rangle^3_X\,|\text{“}fail\text{”}\rangle_Z\,|{-1/2}\rangle^3_Y \ + \\ + \frac{1}{\sqrt{2}}(|fail\rangle^3_X\,|\text{“}fail\text{”}\rangle_Z - |OK\rangle^3_X\,|\text{“}OK\text{”}\rangle_Z)\,|{+1/2}\rangle^3_Y \end{bmatrix} \\
&= \frac{1}{\sqrt{24}}\begin{bmatrix} |heads\rangle^3_X \begin{Bmatrix} \left(3\,|fail\rangle^3_Y + |OK\rangle^3_Y\right)|\text{“}fail\text{”}\rangle_Z \\ + \left(|OK\rangle^3_Y - |fail\rangle^3_Y\right)|\text{“}OK\text{”}\rangle_Z \end{Bmatrix} \\ + \,|tails\rangle^3_X \begin{Bmatrix} \left(3\,|fail\rangle^3_Y + |OK\rangle^3_Y\right)|\text{“}fail\text{”}\rangle_Z \\ - \left(|OK\rangle^3_Y - |fail\rangle^3_Y\right)|\text{“}OK\text{”}\rangle_Z \end{Bmatrix} \end{bmatrix}
\end{aligned} \tag{18}$$

What can Wigner legitimately infer about Xena's outcome at $t = 0$? Prior to $t = 4$ he has yet to perform his own measurement of w on Y, and prior to $t = 5$ he remains unaware of the outcome of Zeus's measurement of z on X at $t = 3$. But even before $t = 4$ Wigner can still use $|\psi\rangle^3_{XYZ}$ to reason hypothetically about Xena's outcome, conditional on Zeus's and his own measurements both having the outcome "OK": on learning at $t = 5$ that these antecedents are true, he can then infer the truth of the consequent of this conditional. Wigner should therefore replace the incorrect reasoning of step 4* as follows.

Step 4 Assume Zeus's measurement of z on X at $t = 3$ had a unique, physical outcome and that there are then no interactions among X, Y, Z prior to $t = 4$. Then the state of XYZ *at* $t = 4$ is

$$|\psi\rangle^4_{XYZ} = \frac{1}{\sqrt{24}}\begin{bmatrix} |heads\rangle^4_X \begin{Bmatrix} \left(3\,|fail\rangle^4_Y + |OK\rangle^4_Y\right)|\text{“}fail\text{”}\rangle^4_Z \\ + \left(|OK\rangle^4_Y - |fail\rangle^4_Y\right)|\text{“}OK\text{”}\rangle^4_Z \end{Bmatrix} \\ + \,|tails\rangle^4_X \begin{Bmatrix} \left(3\,|fail\rangle^4_Y + |OK\rangle^4_Y\right)|\text{“}fail\text{”}\rangle^4_Z \\ - \left(|OK\rangle^4_Y - |fail\rangle^4_Y\right)|\text{“}OK\text{”}\rangle^4_Z \end{Bmatrix} \end{bmatrix} \tag{19}$$

Suppose that Wigner's unique physical outcome on measuring Y at $t = 4$ were "OK". Now consider the hypothesis that Xena's outcome at $t = 0$ was "tails". Equation (19) then implies that the probability of Wigner's outcome "OK" would have been $1/6$. On the alternative hypothesis that Xena's outcome at $t = 0$ was "heads", Eq. (19) also implies that the probability of Wigner getting outcome "OK" would have been $1/6$. So if Wigner were to get outcome "OK" for his measurement at $t = 4$

his knowledge of this outcome would not entitle him to infer the outcome of Xena's measurement at $t = 0$. Indeed, application of Bayes's theorem would lead him to conclude that knowledge of the outcome of his measurement at $t = 4$ should have no effect on his estimate of the probabilities of Xena's possible outcomes: they remain $prob$("heads") $= 1/3$, $prob$("tails") $= 2/3$ after conditionalizing on either possible outcome of his measurement at $t = 4$, and again after further conditionalizing on either possible outcome of Zeus's measurement at $t = 3$.

Consider, for purposes of contrast, how Wigner should reason if he knew that Zeus performed no measurement at $t = 3$. In that case he should assign the following state to XY at $t = 4$:

$$|\psi\rangle^4_{XY} = \frac{1}{\sqrt{3}}(|heads\rangle^4_X\,|{-1/2}\rangle^4_Y + \sqrt{2}\,|tails\rangle^4_X\,|fail\rangle^4_Y). \tag{20}$$

Knowledge of the outcome "OK" of his own measurement of w at $t = 4$ would then entitle him to conclude (with probability 1) that the outcome of Xena's measurement of f on c at $t = 0$ was "heads". This conclusion follows by an inference that parallels step 4* of the reasoning discussed previously. Unlike step 4* itself, the parallel inference is valid because of the assumption that Zeus performed no intervening measurement.

But that same assumption invalidates the premise of step 1 of the reasoning discussed previously. Failing the conclusion of step 1, Wigner would no longer be entitled to take steps 2 and 3. So if he knew that Zeus performed no measurement at $t = 3$ then Wigner could no longer validly conclude that the outcome of Xena's measurement of f on c at $t = 0$ was "tails". In this contrasting case, Wigner should correctly, and consistently, conclude that the unique outcome of Xena's measurement of f on c at $t = 0$ was "heads".

The preceding analysis of the two contrasting cases (with, and without, Zeus's intervening measurement) shows clearly why *Intervention Insensitivity* must be rejected, as inconsistent with *Universality* and *No collapse*. But it may appear to raise a worry about locality. For how can a physically isolated intervening event like Zeus's distant measurement on X have such an impact on Wigner's reasoning about local matters in this scenario?

The form of the question suggests an answer to the worry it seeks to express. Zeus's measurement on X certainly does not influence the outcome of Xena's measurement: if it did, the influence would not be non-local but time-reversed, since Zeus's measurement occurred later than Xena's! Xena's outcome is what it is, irrespective of Zeus's measurement. If Zeus's measurement were to influence anything it would be Wigner's outcome, not Xena's. But Wigner's outcome "OK" has the same probability (1/6) whether or not Zeus performs his measurement. It is only the correlation between Xena's and Wigner's outcomes that differs between the cases where Zeus does and does not measure z on X.

While Zeus's measurement modifies this correlation, it does so despite being causally unrelated to any of its constituent events. This intervention sensitivity is not an instance of non-local causal influence. The suspicion that it is may arise from the view that correlations in non-separable states like $|\psi\rangle^4_{XY}$ are causal because they specify probabilistic counterfactual dependence between the outcomes of distant measurements in violation of Bell inequalities [10]. While controversy continues [13] as to whether such counterfactual dependence constitutes or evidences non-local influence, there are well-known strategies for denying that it does.[4] So the failure of *Intervention Insensitivity* raises no *new* worry about non-locality.

4 A Third Argument

I first heard this argument in a talk by Pusey [15], who there credits it to Luis Masanes. As in sects. 2 and 3 my purpose is not critical but constructive. To understand the limits of quantum objectivity I look to develop the strongest arguments against the assumption that every quantum measurement can have a definite, objective outcome.

Once again, the argument is set in the context of a *Gedankenexperiment* featuring four experimenters. For variety I have changed their names to Alice, Bob, Carol and Dan. But while Carol and Dan perform difficult but technically feasible lab experiments, Alice and Bob are credited with even more extreme abilities than the Zeus and Wigner who figured in the previous arguments (though their exercise of these powers involves no violation of unitary quantum theory).

Each of Alice and Bob are in their own separate laboratories, totally physically isolated except for a shared Bohm–EPR pair of spin-$1/2$ particles on which they intend to perform measurements of (normalized) spin-components, one on each particle from the pair. Alice is to measure magnitude A_a corresponding to operator $\hat{A}_a$ with eigenvalues $\{+1, -1\}$ on particle 1, while Bob is to measure magnitude B_b corresponding to operator $\hat{B}_b$ with eigenvalues $\{+1, -1\}$ on particle 2. a, b label two directions in space along which Alice and Bob (respectively) set the axes of their spin-measuring devices. Alice will choose setting a and perform measurement of A_a at spacelike-separation from Bob's choice of setting b and measurement of B_b.

But before performing these measurements, Alice and Bob first delegate a similar task to their friends, Carol and Dan respectively. Carol occupies her own separate laboratory, initially totally physically isolated from Alice's: Dan occupies his own separate laboratory, initially totally physically isolated from Bob's. Carol and Dan perform measurements on the Bohm–EPR pair: Carol measures (normalized) spin-component C_c on particle 1, while Dan measures (normalized) spin-component D_d on particle 2. Carol's choice of setting c and measurement of C_c are each spacelike-separated from Dan's choice of setting d and measurement of D_d. Assume Carol's and Dan's measurements each have a definite, physical outcome that is registered, recorded and experienced by them separately in their labs.

It is important to note that Alice and Carol both perform their measurements on the very same particle 1, and that Bob and Dan perform their measurements on the very same particle 2. To make this possible, after performing Carol's measurement particle 1 is transferred out of her lab and into Alice's lab, and after performing Dan's measurement particle 2 is transferred out of his lab and into Bob's lab. Assume that measurement causes no physical "collapse" of the quantum state, so that each spin-measurement proceeds in accordance with a unitary interaction between the measured

[4] My preferred strategy [14] depends on an interventionist approach to causal influence.

particle and the rest of the experimenter's lab, and that this is consistent with its having a definite, physical outcome recorded by the experimenter in that lab. It follows that Carol's measurement entangles the state of her lab C with that of particle 1, while Dan's measurement entangles the state of his lab D with that of particle 2.

But Alice and Bob use their superpowers to undo these entanglements by applying very carefully tailored interactions, in the first case between 1 and C, and in the second case between 2 and D. This restores C and D to their original states, and also restores the original spin-entangled state of $1 + 2$. That is how it is possible for Alice and Bob to perform spin-measurements on the same Bohm–EPR pair as Carol and Dan.

By assumption, we now have a situation in which successive measurements of spin-component (in the c and a directions) have been performed on particle 1 of an individual Bohm–EPR pair, while successive measurements of spin-component (in the d and b directions) have been performed on particle 2 of that pair. By assumption, each of these measurements has a definite, physical outcome registered, recorded and experienced by an experimenter in his or her laboratory. Finally suppose that this entire situation is repeated very many times, each time with a different Bohm–EPR pair, giving rise to a statistical distribution of results for the four outcomes in each trial.

We may use quantum theory to predict the corresponding probability distribution by applying the Born rule to appropriate quantum states. From Alice's perspective, events in a given trial unfold in the following sequence. At time t_0 the particles are in state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle_1 |\downarrow\rangle_2 - |\downarrow\rangle_1 |\uparrow\rangle_2) \tag{21}$$

while C, D are in states $|ready\rangle_C$, $|ready\rangle_D$ respectively. Then Dan measures the d-spin of 2 by means of a unitary interaction $\hat{U}_D^2$ as follows

$$\begin{aligned} \hat{U}_D^2 |\downarrow_d\rangle_2 |\text{ready}\rangle_D &= |d\text{-down}\rangle_{2D} \\ \hat{U}_D^2 |\uparrow_d\rangle_2 |\text{ready}\rangle_D &= |d\text{-up}\rangle_{2D} . \end{aligned} \tag{22}$$

So at time t_1 when Dan has recorded the definite outcome as either d-down or d-up, Alice assigns the following state to $12D$

$$\Psi_1 = \frac{1}{\sqrt{2}}(|\uparrow_d\rangle_1 |d\text{-down}\rangle_{2D} - |\downarrow_d\rangle_1 |d\text{-up}\rangle_{2D}). \tag{23}$$

Shortly after t_1 Carol measures the c-spin of particle 1 by a unitary interaction $\hat{U}_C^1$

$$\begin{aligned} \hat{U}_C^1 |\downarrow_c\rangle_1 |\text{ready}\rangle_C &= |c\text{-down}\rangle_{1C} \\ \hat{U}_C^1 |\uparrow_c\rangle_1 |\text{ready}\rangle_C &= |c\text{-up}\rangle_{1C} . \end{aligned} \tag{24}$$

So at time t_2 when Carol has recorded the definite outcome as either c-down or c-up, Alice assigns the following state to $12DC$

$$\Psi_2 = \frac{1}{\sqrt{2}}(\hat{U}_C^1 \left|\uparrow_d\right\rangle_1 \left|\text{ready}\right\rangle_C \left|d\text{-down}\right\rangle_{2D} - \hat{U}_C^1 \left|\downarrow_d\right\rangle_1 \left|\text{ready}\right\rangle_C \left|d\text{-up}\right\rangle_{2D}) \quad (25)$$

Next Alice "undoes" the effects of Carol's measurement by applying an interaction between 1 and C with unitary $\hat{U}_C^{1\dagger}$, and assigns the state Ψ_3 at time t_3 to $12D$ (which is no longer entangled with that of C)

$$\Psi_3 = \frac{1}{\sqrt{2}}(\left|\uparrow_d\right\rangle_1 \left|d\text{-down}\right\rangle_{2D} - \left|\downarrow_d\right\rangle_1 \left|d\text{-up}\right\rangle_{2D}). \quad (26)$$

Shortly after t_3, Alice measures a-spin on 1 and at time t_4 gets a definite, physical outcome of either a-down or a-up. Then Bob "undoes" Dan's measurement on particle 2 by implementing an interaction in accordance with unitary $\hat{U}_D^{2\dagger}$, before measuring the b-spin of 2 and at time t_5 getting a definite, physical outcome of either b-down or b-up.

By applying the Born rule to the Bohm–EPR spin-state at t_0, Alice predicts the probabilistic correlation function $E(c, d)$ for Carol's and Dan's measurement outcomes

$$E(c, d) = -\cos(c - d). \quad (27)$$

To predict the correlation function $E(a, d)$ for Alice's and Dan's measurement outcomes, Alice reasons as follows. If Carol had performed no measurement and C and 1 had never interacted, then between t_1 and t_4 Alice and Dan would just have been recording a correlation between outcomes of an a-spin measurement on 1 and an earlier d-spin measurement on 2—a standard Bell experiment with settings and measurements performed at timelike separation. For such a case, the Born rule predicts

$$E(a, d) = -\cos(a - d). \quad (28)$$

In the present case, C and 1 interacted twice between t_1 and t_4, but these interactions had no overall effect on the state of the joint system $12D$ at the time when Alice performed her measurement of a-spin: its state was the same at t_3 as it had been at t_1 ($\Psi_3 = \Psi_1$). It follows that in the present case also quantum theory predicts the correlation function

$$E(a, d) = -\cos(a - d). \quad (29)$$

After the effects of Dan's measurement on 2 have been "undone" by Bob's implementation of the interaction $\hat{U}_D^{2\dagger}$, Alice should again recognize that the outcomes of her measurement of a-spin on 1 and Bob's spacelike-separated measurement of b-spin on 2 constitute a record of a correlation in a standard (spacelike separated) Bell experiment, with predicted correlation function

$$E(a, b) = -\cos(a - b). \quad (30)$$

So far we have been considering the events involved in a single trial from Alice's perspective. But those same events should also be considered from the perspective of Bob. If the labs of Alice, Bob and friends are all in the same state of motion, then the events we have considered will play out in the same sequence also from Bob's perspective. But it is well known that the time-order of spacelike separated events is not invariant under transformations of inertial frame.

Suppose that Alice's lab and Carol's lab are in one state of motion relative to frame F (moving to the right at speed v, say), while Bob's lab and Carol's lab are in a different sate of motion (moving to the left at speed v, say). To make sure that Alice is in position to manipulate 1 and C we may assume that they both remain inside, and move together with, Alice's lab A: and to make sure that Bob is in position to manipulate 2 and D we may assume that they both remain inside, and move together with, Bob's lab B. This arrangement is depicted in Fig. 1. Relative to the state of motion of Bob and Dan, the same events play out over a period marked by the sequence of times $\langle t_0^*, t_1^*, t_2^*, t_3^*, t_4^*, t_5^* \rangle$.

Note that in the *'d frame Carol's measurement precedes Dan's and Bob's precedes Alice's. Most important, note that the state of $12C$ is the same at t_3^* as at t_1^*. Paralleling Alice's reasoning, Bob should therefore conclude that in this situation the correlation function for his outcome when measuring the b-spin of 2 and Carol's outcome when measuring the c-spin of 1 is $E(b, c) = -\cos(b - c)$.

It is a central assumption of this third argument that every spin measurement by Alice, Bob, Carol or Dan has a definite, physical outcome—either spin up or spin down with respect to the chosen direction. It follows that in a long sequence of trials of the *Gedankenexperiment* just described there will be a statistical distribution of actual outcomes, with a set of outcomes that may be labeled $\langle a, b, c, d \rangle$ in each trial. Statistical correlations between pairs of actual experimental outcomes may be represented in the usual way by statistical correlation functions $corr(a, b)$, $corr(b, c)$, $corr(c, d)$, $corr(a, d)$. It follows that these statistical correlations will satisfy the inequality

$$|corr(a, b) + corr(b, c) + corr(c, d) - corr(a, d)| \leq 2. \tag{31}$$

Note that no locality assumption is required to derive this inequality here, since it is mathematically equivalent to the existence of a joint distribution over the actual, physical outcomes whose existence has been assumed [16].

But we saw that quantum mechanics predicts probabilistic correlation functions $E(a, b)$, $E(b, c)$, $E(c, d)$, $E(a, d)$ for these pairs of outcomes that may be compared to the inequality

$$|E(a, b) + E(b, c) + E(c, d) - E(a, d)| \leq 2. \tag{32}$$

It is well known that quantum theory predicts violation of inequality (32) for certain choices of directions a, b, c, d. If the particles and labs of the experimenters in the *Gedankenexperiment* had been at relative rest, then the choice of four directions in a plane defined by rotations of $a^\circ = 0^\circ$, $b^\circ = 45^\circ$, $c^\circ = 90^\circ$, $d^\circ = 135^\circ$ from a fixed

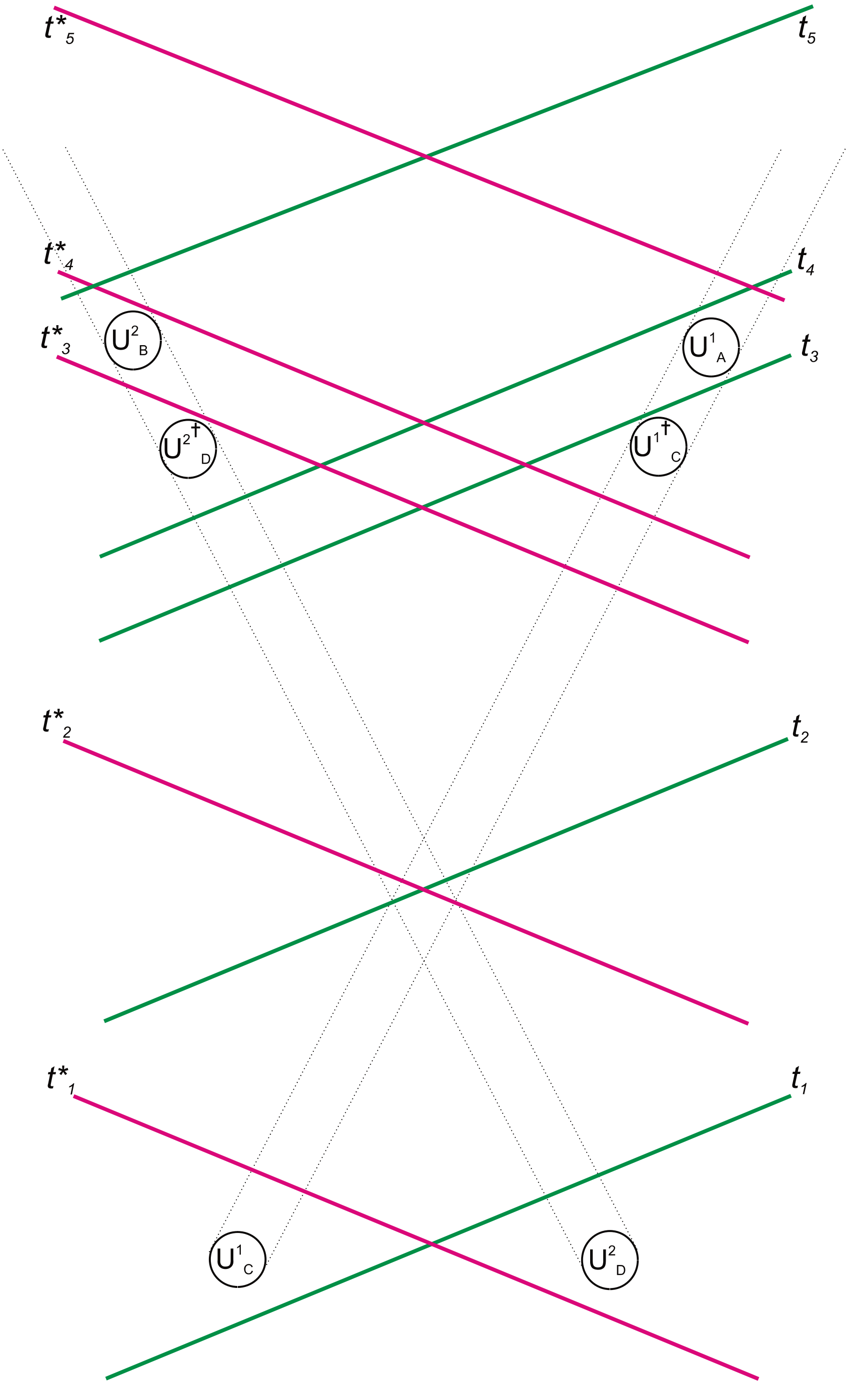

Fig. 1 Spacetime diagram for third argument

axis would yield maximal violation of (32) with predicted value

$$|E(a,b)+E(b,c)+E(c,d)-E(a,d)|=2\sqrt{2}. \tag{33}$$

The relativistic relative motion of labs and particles makes it necessary to take account of the associated Wigner rotation of vectors, affecting the predicted value for this choice of directions. But the inequality is still maximally violated for a different choice of directions.[5]

5 Conclusion

Each of the three arguments analyzed in this paper sought to establish a contradiction between the universal applicability of unitary quantum theory and the assumption that a well-conducted quantum measurement always has a definite, physical outcome. The first argument succeeded in doing so only by implicitly relying on assumptions that the work of Bell [10] and Kochen and Specker [18] gives us good reasons to reject—in Einstein's [19] words, that in the circumstances described in the associated *Gedankenexperiment* the individual system (before the measurement) has a definite value for all variables of the system, and more specifically, *that* value which is determined by a measurement of this variable. Failing some such naive realist assumptions, nothing justifies the argument's application of quantum theory to predict probabilities for outcomes of hypothetical measurements which would be incompatible with those actually performed.

Though it does not rely on such naive realist assumptions, the second argument also depends on a superficially plausible assumption about the outcomes of counterfactual measurements I called *intervention insensitivity*, according to which the truth-value of an outcome-counterfactual is insensitive to the occurrence of a physically isolated intervening event. But in the circumstances of the associated *Gedankenexperiment*, intervention insensitivity is itself incompatible with the universal applicability of unitary quantum theory. Since a contradiction then follows even if each (well-conducted) quantum measurement does *not* have a definite, physical outcome in the *Gedankenexperiment*, the argument does not establish its intended conclusion.

Unlike the first two arguments, the third argument relies on no implicit assumptions about the outcomes of hypothetical measurements, since all the outcomes it considers are assumed to be actual. I think it succeeds in showing that, in the circumstances described in the associated *Gedankenexperiment*, the universal applicability of unitary quantum theory implies (with probability approaching 1) that there is no consistent assignment of values to the (supposedly definite, physical) outcomes of the measurements in the sequence of trials there considered.

This result prompts further reflection on how to understand quantum theory. But the circumstances of the *Gedankenexperiment* in the third argument are so extreme as forever to resist experimental realization. There are no foreseeable circumstances in

5 [17] specifies the necessary directions in Sect. 4. Rather than being coplanar (with respect to frame F) these may be chosen to lie on a cone centered on the direction of motion of the lab in which that spin measurement is performed.

which the argument would require us to deny that a well-conducted quantum measurement has a definite, physical outcome. The arguments considered in this paper give us no reason to doubt the sincerity or truth of experimenters' reports of definite, physical outcomes. But I think the third argument should make us reconsider the extent and nature of their objectivity. This paper was intended to both motivate and prepare the way for the pursuit of that project.

Acknowledgements Thanks to Jeff Bub for a helpful correspondence on Frauchiger and Renner's argument, to Časlav Brukner for conversations and correspondence over several years, and to a reviewer for good strategic advice. None of them should be taken to endorse the analysis or conclusions of this paper.

Appendix

In restating the argument of [9,12] I have changed the notation to try to make it easier to follow. The following table supplies a translation between my notation and that used in [12].

Agent	Lab	Measured system	Measured observable	Other observable
$\bar{F} \leftrightsquigarrow$ Xena	$\overline{L} \leftrightsquigarrow X$	$R \leftrightsquigarrow c$	heads/tails $\leftrightsquigarrow f$	
$F \leftrightsquigarrow$ Yvonne	$L \leftrightsquigarrow Y$	$S \leftrightsquigarrow s$	up/down $\leftrightsquigarrow S_z$	
$\bar{W} \leftrightsquigarrow$ Zeus	$\bullet \leftrightsquigarrow Z$	$\bar{L} \leftrightsquigarrow X$	$\bar{w} \leftrightsquigarrow z$	$\bullet \leftrightsquigarrow x$
$W \leftrightsquigarrow$ Wigner	$\bullet \leftrightsquigarrow W$	$L \leftrightsquigarrow Y$	$w \leftrightsquigarrow w$	$\bullet \leftrightsquigarrow y$

Readers familiar with Wigner's original "friend" argument [11] will be primed to attribute extraordinary powers to the experimenter I have named Wigner, and I thought it appropriate to name a second character with such almost "God-like" powers Zeus. This naturally suggested also giving the experimenters charged with less extraordinary tasks names whose initial letters are also at the end of the alphabet, with corresponding labels for their labs and measured observables.

While such changes are merely cosmetic, my restatement deliberately lacks one feature emphasized by the authors of the argument of [12] that they call "consistent reasoning", illustrate in their Fig. 1, and formalize in their assumption (C). Both in the original and in my restatement it is Wigner (W) whose reasoning is the ultimate focus of the argument. But the authors of the original argument consider it important that Wigner's reasoning incorporates the reasoning of the other experimenters [*via* assumption (C)].

It is vital to check whether Wigner's reasoning is both internally consistent and consistent with the reasoning of the other experimenters in this *Gedankenexperiment*. My restatement makes it clear how Wigner can consistently apply quantum theory without considering the reasoning of any other experimenters. But are the conclusions of this independent reasoning by Wigner consistent with those of the other experimenters, based on their own applications of quantum theory? Indeed they are, provided each experimenter has applied quantum theory *correctly*. The problem with the argument of Frauchiger and Renner is that one experimenter (Xena/$\bar{F}$) has applied quantum theory *incorrectly*.

Recall step 4* of the reasoning in my restatement of this argument (see §3). I attributed this reasoning to Wigner, while pointing out that Zeus's subsequent measurement of z renders it fallacious. Frauchiger and Renner initially attribute parallel reasoning to Xena/$\bar{F}$ and then use assumption (C) to attribute its conclusion also to Wigner. To see where things go wrong if Xena/$\bar{F}$ reasons this way, I quote from [12].

"Specifically, agent $\bar{F}$ may start her reasoning with the two statements

$$s_I^{\bar{F}} = \text{"If } r = tails \text{ at time } n : 10 \text{ then spin } S \text{ is in state } |\rightarrow\rangle_S \text{ at time } n : 10\text{"}$$
$$s_M^{\bar{F}} = \text{ The value } w \text{ is obtained by a measurement of } L \text{ w.r.t.} \{\pi_{ok}^H, \pi_{fail}^H\}\text{".''}$$

They conclude that $\bar{F}$ can infer from $s_I^{\bar{F}}$ and $s_M^{\bar{F}}$ that statement $s_Q^{\bar{F}}$ holds:

$$s_Q^{\bar{F}} = \text{"If } r = tails \text{ at time } n : 10 \text{ then I am certain that } W \text{ will observe } w = fail \text{ at } n : 40\text{".}$$

Starting with $s_Q^{\bar{F}}$, they then apply assumption (C) to the reasoning of the other agents successively, eventually to establish that Wigner may conclude

$$s_2^W = \text{" If } \bar{w} = \overline{ok} \text{ at time } n : 30 \text{ then I am certain that I will observe } w = fail \text{ at } n : 40\text{",}$$

which (given (S)) is inconsistent with W's independent conclusion (based on assumption (Q))

s_Q^W ="I am certain that there exists a round $n \in \mathbb{N}_{\geq 0}$ in which it is announced that $\bar{w} = \overline{ok}$ at time $n : 30$ and $w = ok$ at $n : 40$."

But this chain of reasoning is based on a mistaken starting point, since $\bar{F}$ has applied quantum theory incorrectly in asserting statement $s_Q^{\bar{F}}$. Compare $s_Q^{\bar{F}}$ with the corresponding conclusion of Wigner's fallacious reasoning in step 4* of §3:

"If the unique outcome of Xena's measurement of f on c at $t = 0$ had been "tails", the unique outcome of my measurement of w on Y at $t = 4$ would have been "fail".

Agent $\bar{F}$'s reasoning was equally fallacious here. The problem starts with statement $s_I^{\bar{F}}$: $\bar{F}$ is correct to assign state $|\rightarrow\rangle_S$ to S at time $n : 10$ for certain purposes but not for others. Suppose, for example, that $\bar{F}$ had "flipped the quantum coin R" by passing that system through the poles of a Stern–Gerlach magnet. By applying unitary quantum theory, $\bar{F}$ should conclude that this will induce no physical collapse of R's spin state but entangle it with its translational state, and thence with the rest of her lab [20]. So while $\bar{F}$ would be correct then to assign state $|\rightarrow\rangle_S$ to S at time $n : 10$ for the purpose of predicting the outcome of a subsequent spin measurement on S alone, she would be incorrect to assign state $|{\rightarrow}\rangle_S$ to S at time $n:10$ for the purpose of predicting correlations between S (or anything with which it subsequently interacts) and her lab $\bar{L}$ (or anything with which it subsequently interacts).

By using the phrase 'is in', statement $s_I^{\bar{F}}$ ignores the essential relativity of S's state assignment at time $n:10$ to these different applications. By using $s_I^{\bar{F}}$ to infer $s_M^{\bar{F}}$, agent $\bar{F}$ is, in effect, taking $\bar{F}$'s coin flip to involve the physical collapse of R's state rather than the unitary evolution represented by Eq. (8). So agent $\bar{F}$ is mistaken to assert $s_Q^{\bar{F}}$, and W would be wrong to incorporate this mistake in his own reasoning by applying assumption (C).

Frauchiger and Renner [12] justify $\bar{F}$'s inference from $s_I^{\bar{F}}$ and $s_M^{\bar{F}}$ to $s_Q^{\bar{F}}$ by appeal to assumption (Q). I have argued that $\bar{F}$ is not justified in asserting $s_Q^{\bar{F}}$, since $\bar{F}$ is justified in using the state assignment licensed by $s_I^{\bar{F}}$ for the purpose of predicting the outcome of a measurement on S only where S's correlations with other systems (encoded in an entangled state of a supersystem) may be neglected. But the sequence of interactions in the *Gedankenexperiment* successively entangle the state of S with those of R, $\bar{L}$, L and $\bar{W}$. So in reasoning about the outcome of W's measurement of w, $\bar{F}$ must take account of this progressive entanglement of the states of S and $\bar{W}$.

Specifically, to predict the outcome of W's measurement of w, $\bar{F}$ must represent that measurement as the second part of W's joint measurement on the system $\bar{W}+L$. This interaction between W and $\bar{W}$ was represented in §3 as the apparently innocuous Step 1 in which Wigner simply asked Zeus what was the outcome of his measurement. But it is not this interaction but the prior interaction between $\bar{W}$ and L that undercuts $\bar{F}$'s justification for using the state assignment $|{\rightarrow}\rangle_S$ in inferring $s_Q^{\bar{F}}$ from $s_I^{\bar{F}}$ and $s_M^{\bar{F}}$. Only by neglecting the prior interaction between $\bar{W}$ and L can $\bar{F}$ draw the erroneous conclusion $s_Q^{\bar{F}}$.

*W*igner can reason consistently about the unique, physical outcomes of all experiments in the *Gedankenexperiment* of [9,12] without any appeal to the reasoning of the other agents involved. Each of these other agents may reason equally consistently. And their collective reasoning is perfectly in accord with assumption (C) as well as the universal applicability of unitary quantum theory and the existence of a unique, physical outcome of every measurement that figures in the *Gedankenexperiment* of [9,12].

References

1. Fuchs, C.: QBism, the perimeter of quantum bayesianism (2010). arXiv:1003.5209
2. Fuchs, C., Mermin, N.D., Schack, R.: An introduction to QBism with an application to the locality of quantum mechanics. Am. J. Phys. **82**, 749–54 (2014)
3. Brukner, Č.: On the quantum measurement problem. In: Bertlmann, R., Zeilinger, A. (eds.) Quantum Unspeakables II, pp. 95–117. Springer International, Cham (2017)
4. Brukner, Č.: A no-go theorem for observer-independent facts. Entropy **20**, 350 (2018)
5. Rovelli, C.: Relational quantum mechanics. Int. J. Theor. Phys. **35**, 1637–78 (1996)
6. Popper, K.R.: The Logic of Scientific Discovery. Hutchinson, London (1959)
7. Deutsch, D.: Quantum theory as a universal physical theory. Int. J. Theor. Phys. **24**, 1–41 (1985)

8. Wallace, D.: The Emergent Multiverse. Oxford University Press, Oxford (2012)
9. Frauchiger, D., Renner, R.: Single-world interpretations of quantum theory cannot be self-consistent" (2016). https://arxiv.org/abs/1604.07422
10. Bell, J.S.: Speakable and Unspeakable in Quantum Mechanics, Revised edn. Cambridge University Press, Cambridge (2004)
11. Wigner, E.: Remarks on the mind-body question. In: Good, I. (ed.) The Scientist Speculates. Heinemann, London (1961)
12. Frauchiger, D., Renner, R.: Quantum theory cannot consistently describe the use of itself. Nat Commun. **9**, 3711 (2018)
13. Maudlin, T.: Quantum Non-locality and Relativity, 3rd edn. Wiley-Blackwell, Chichester (2011)
14. Healey, R.A.: Locality, probability and causality. In: Bell, M., Shan, G. (eds.) Quantum Nonlocality and Reality, pp. 172–94. Cambridge University Press, Cambridge (2016)
15. Pusey, M.:"Is QBism 80% complete, or 20%", talk given at the *Information-Theoretic Interpretations of Quantum Mechanics* workshop, Western University, Canada. Available at https://grfilms.net/v-matthew-pusey-is-qbism-80-complete-or-20-_9Rs61l8MyY.html) (2016)
16. Fine, A.: Joint distributions, quantum correlations, and commuting observables. J. Math. Phys. **23**, 1306–10 (1982)
17. Lee, D., Chang-Young, E.: Quantum entanglement under Lorentz boost. New J. Phys. **6**, 67 (2004)
18. Kochen, S., Specker, E.: The problem of hidden variables in quantum mechanics. J. Math. Mech. **17**, 59–87 (1967)
19. Einstein, A.: Autobiographical notes. In: Schilpp, P. (ed.) Albert Einstein: Philosopher-Scientist. Open Court Press, La Salle (1949)
20. Wigner, E.: The problem of measurement. Am. J. Phys. **31**, 6–15 (1963)