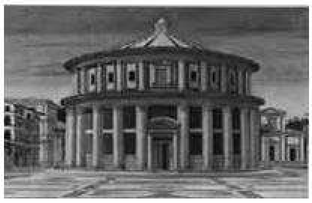
2022 Isonomia – Epistemologica
Rivista online di Filosofia
Università degli Studi di Urbino Carlo Bo



Wigner's Friend Extended Thought Experiment
A Philosophers' look

Alberto Corti[1,2], Vincenzo Fano[2] & Gino Tarozzi[2]
[1]University of Geneva; [2]University of Urbino
a.corti1@campus.uniurb.it; vincenzo.fano@uniurb.it; gino.tarozzi@uniurb.it

Abstract. Frauchiger and Renner (2018) presented a no-go theorem inspired by the well-known Wigner's friend paradox. Their work is a remarkable contribution to the foundations of quantum mechanics, insofar as it allows evaluation of the interpretations of the theory based on the assumptions that must be dropped to circumvent the contradiction. This paper aims to review the philosophically salient aspects of the paradox. To do so, we firstly introduce the original Wigner's paradox and its consequences. Then we make explicit the logical structure of the extended paradox through the aid of multi-agent epistemic logic. Such work is useful not only for presenting a more accessible formulation of the paradox but also for evaluating its bearings on the main interpretations of quantum mechanics. We conclude that despite the fact that prima facie the paradox encourages an antirealist view, some possibilities are left open for realist interpretations of quantum mechanics.

Keywords. Quantum mechanics, Wigner's friend paradox, Foundations of physics, Epistemology of Science, No-go theorem.

1 Introduction

Frauchiger and Renner (2018) recently presented a reformulation of the thought experiment proposed by Wigner (1961). Their paper started a hot debate in the field of foundations of physics, addressed mainly by physicists and logicians (Bub, (2017); Sudbery, (2017); Fortin and Lombardi, (2019); Lazarovici and Hubert, (2019)).

A. Corti, V. Fano, G. Tarozzi,
"Wigner's Friend Extended Thought Experiment. A Philosophers' look",

University of Urbino Carlo Bo
http://isonomia.uniurb.it/epistemologica

Their formulation of the paradox shows many outstanding features. For instance, a contradiction is reached assuming just three epistemological principles, in contrast to the original formulation, which leads to a peculiar metaphysical thesis rather than to a real inconsistency. Furthermore, since the assumptions are so few, the paradox is extremely solid: there are not many possibilities available to bypass it. Moreover, since the principles mentioned above are, as we will see, quite reasonable and logically well-grounded, the denial of one of them brings a heavy (epistemological) price.
As is the case for other no-go theorems, this is a profound and meaningful result on its own. Indeed, it allows us to compare the interpretations of non-relativistic quantum mechanics (QM hereafter) based on which principle must be dropped to avoid the reached contradiction. It is a well-known fact that the interpretations of QM do not differ as much in their empirical predictions as they do in their metaphysical description of the world. Precisely because the clash between interpretations takes place mostly in the metaphysical and epistemological arena, the results obtained by no-go theorems have great importance. Indeed, the metaphysical pictures drawn by quantum interpretations are already quite unusual, if compared to the classical image to which we are accustomed. Therefore, being committed to the denial of new reasonable principles has vast consequences on the evaluation of the interpretations of QM. Another remarkable feature of Frauchinger and Renner's paradox is that the assumptions are not strictly speaking physical constraints, but epistemological principles. Their epistemological character led us to formulate the argument within a formal epistemology framework. Specifically, in this paper, we reframe the paradox with the aid of multi-agent epistemic logic (Fagin, Moses, Halpern & Vardi 2003). This setting not only allows us to pinpoint the logical structure of the argument but also to show distinctly the price paid to abandon one of the principles mentioned above. Moreover, it is useful in highlighting logical connections assumed implicitly as well as results that are unexpected and could otherwise remain unnoticed.
In the first section (§2) we present the original formulation by Wigner and some of its implications with respect to foundational issues; in particular, we argue that the paradox does not run against the main realist interpretations of QM. In the next section, we present Frauchinger and Renner's reformulation of the paradox. We restate (§3.1) formally the principles assumed by the authors before presenting a simplified, yet formal, rewriting of the contradiction reached in the paradox (§3.2). The simplification is two-fold. On the one hand, we omitted some details of the paradox, supposing that our conclusions do

not hinge on them.[1] On the other, to reach a wider audience, we chose to be less rigorous than others with the logic employed. Nonetheless, reframing the paradox in a logical form is necessary for assessing the consequences that the contradiction proved by Frauchinger and Renner has on the main interpretations of QM; these implications are discussed in (§4). A conclusion (§5) follows the evaluation.
We must acknowledge that as we have been finishing our draft of this paper, papers similar to ours have been published. The most similar is the work of Nurgalieva and del Rio (2019), in which they carefully reconstruct Frauchinger and Renner's paradox to argue that epistemic logic falls short in QM contexts. The spirit of our paper, though, is quite different: our aim is to give a more accessible and well-rounded introduction to the paradox. Our discussion is more general, with a focus on how a family of epistemic logics can be applied to the context of the extended paradox and what options are left open for realist interpretations of QM. Where Nurgalieva and del Rio, like Boge (2019), offer a more detailed and logically rigorous formalization of the thought experiment, we focus more on its bare structure and its implications for the interpretations of QM. Despite the similarities then, the spirit and the contents of our works differ significantly. In summary, Boge transforms Frauchinger and Renner's argument into a theorem in epistemic logic, whereas Nurgalieva and del Rio emphasize the inadequacy of epistemic logic in the quantum context. We, in contrast, assume that such logic holds good in order to evaluate the price that the main interpretations of QM have to pay to deal with the paradox.

2 The original Wigner's paradox

As is well known, the state of a system can be, in respect to an observable (e.g. the spin along the z-axis), in a superposition of the corresponding operator's eigenvectors (in the case at hand, the possible eigenvectors are spin up along the z-axis, $|\uparrow_z\rangle$, and spin down on the same axis, $|\downarrow_z\rangle$): if we measure with a Stern-Gerlach the system's (let us say a silver atom, s hereafter) spin along the z-axis, we will find it by applying the Born rule (**(BR)** henceforth), half the time in the state $|\uparrow_z\rangle_s$ and half the time in state $|\downarrow_z\rangle_s$.[2] Before the measurement,

[1]For instance, we neglected the probability amplitude from many formulas, ignored the time dimension of the thought experiment and omitted the demonstration of statements b) and c); Cf. §3.2.

[2]Throughout the paper, we use subscripts inside a ket to state the observable of which such a state is an eigenstate of (in this case 'z' stands for 'spin along the z-axis'); subscripts outside

s can be described using the following state:

$$|\psi\rangle_s = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle_s + |\downarrow_z\rangle_s) \tag{1}$$

Such a superposed state cannot be understood to be as if the silver atom has either spin up or spin down, nor as if it has both at the same time nor as if it has none. Instead, states like (1) describe something different to the just mentioned classically allowed logical possibilities.[3]

This well-known aspect of QM is single-handedly responsible for most QM paradoxes and classically inconceivable consequences. In what follows, we will focus on the so-called 'Wigner's friend paradox' (Wigner 1961). Even if it is frequently called 'paradox' throughout the literature, it is not. Instead, it is an argument against the possibility of applying QM to systems at all scales and, as Wigner intended, in favour of the causal role of consciousness in the process of wave function collapse.

Wigner's argument is the following (see fig. 1). Let us suppose that s is in the state (1) and that a friend of Wigner, F, measures the spin of the system with a device that we will call a. Furthermore, let us suppose that a can be in three states: ready to perform a measurement, i.e. the apparatus is well-positioned in the lab ($|ready\rangle_a$), measuring that s has spin up ($|\uparrow_z\rangle_a$) and measuring that s has spin down ($|\downarrow_z\rangle_a$). If the system, composed of s and a, follows QM, before the measurement its state will be:

$$|\psi\rangle_{as} = |ready\rangle_a \otimes \frac{1}{\sqrt{2}}(|\uparrow_z\rangle_s + |\downarrow_z\rangle_s) \tag{2}$$

As soon as the silver atom passes through the measurement apparatus the state will evolve according to the principle that governs the dynamics of quantum systems:

(D) Dynamics: The state of a quantum system evolves according to the Schrödinger equation.

a ket are used for specifying of which system the ket represents a state of (in the case at hand, '$|\uparrow_z\rangle_s$', meaning that the system s is in state '$|\uparrow_z\rangle$'.

[3]For an introductory yet precise description of this feature and its consequence see: Albert (1992, Ch. 1).

Hence, the state of the silver atom and the measurement apparatus will evolve – after they have interacted – as follows:

$$|\psi\rangle_{as} = \frac{1}{\sqrt{2}}(|{\uparrow_z}\rangle_a |{\uparrow_z}\rangle_s + |{\downarrow_z}\rangle_a |{\downarrow_z}\rangle_s) \tag{3}$$

That is to say, the states of the silver atom and the Stern-Gerlach will be entangled. Moreover, let us suppose that F performs the measurement in a closed room and that Wigner (W henceforth) is waiting outside, so that he does not know what the experiment's outcome is as seen by F. When F leaves the room, W asks him/her about the experimental result, and he/she answers, let us say, 'spin down'.[4] If this is so, then (3) can no longer be a good physical description of the joint system $a+s$. If F has seen that s had spin down, then the joint system is not in a superposition anymore (i.e. its state cannot be something like (3)), but rather it will have collapsed into:[5]

$$|\psi\rangle_{as} = |{\downarrow_z}\rangle_a |{\downarrow_z}\rangle_s \tag{4}$$

Now, if the room is open, W can check whether the state of the system was $|{\downarrow_z}\rangle_a |{\downarrow_z}\rangle_s$ (for instance, he can check which part of the screen has been hit by the atom, and infer what spin the atom must have had for being deviated so by the Stern-Gerlach), and no problem arises. However, let us suppose further that the room is closed, and W cannot enter (but he can communicate with F, with a phone or other device). Just applying QM's evolution, the state of the joint system $a+s$ should be (3) according to W's perspective; still, F replies that the state, after the measurement, has changed from (3) to (4). W wants to know why the theory's prediction and F's report mismatch, so W then presses F further, asking whether before answering to the previous question he/she already knew the outcome of the measurement. For simplicity, let us assume that F could have been in two possible states, that is to say, he/she could have measured that the system is in state spin up ($|F \uparrow_z\rangle$) or in the state spin down ($|F \downarrow_z\rangle$);[6] then, before asking, F should have been entangled with

[4] i.e. that he/she has found that the silver atom was in the state $|{\downarrow_z}\rangle_s$.

[5] Following the **Projection (or Collapse) Postulate** ((**CP**) henceforth), according to which immediately after a measurement of an observable O (corresponding to some Hermitian operator $\hat{O}$) the state of a system collapses into one of the eigenstates of the corresponding operator.

[6] In what follows, we will write the name of an epistemic agent inside a ket, before the effective state, for pointing out clearly who has measured the state. A state like '$|A \uparrow_z\rangle_s$' will be a shortcut for the state 'measuring system s, epistemic agent A obtained as a result that the

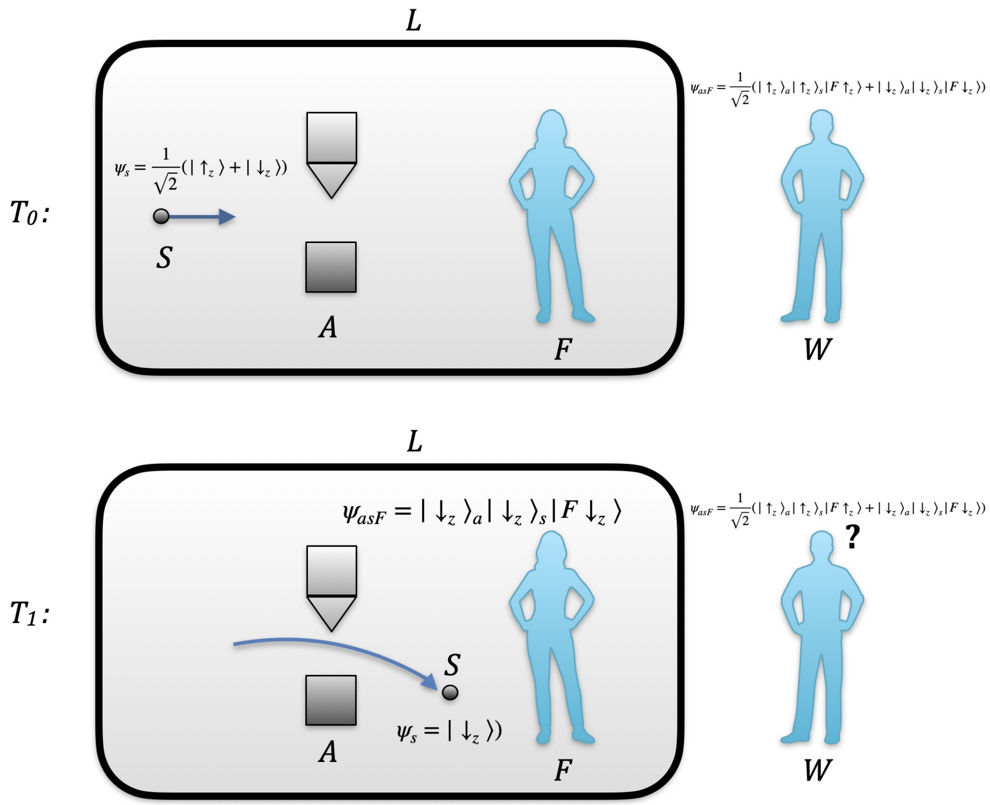


Figure 1: **Wigner's Paradox.** Inside a laboratory L, an agent F measures the spin polarisation along the z-axis of a silver atom s in state (1). After F performs a measurement upon s, the quantum state attributed to s by F is different from that attributed to s by W.

the system $a+s$ and, as a result, F should be superposed between knowing that the outcome is spin up and knowing that it is spin down. Before answering, the joint system of $a+s+F$ should be in a state like:

$$|\psi\rangle_{asF} = \frac{1}{\sqrt{2}}(|\uparrow_z\rangle_a|\uparrow_z\rangle_s|F\uparrow_z\rangle + |\downarrow_z\rangle_a|\downarrow_z\rangle_s|F\downarrow_z\rangle) \tag{5}$$

Wigner (1961)'s hypothesis is that F would instead answer that he/she already knew the result before W's question. The plausibility of such a hypothesis lies in the fact that we never experience in our consciousness any kind of quantum superposition. Therefore, claiming that F would feel otherwise is empirically

system has spin up along the z-axis'.

unjustified. According to F (and contrary to W's prediction), the joint systems $a+s+F$ before his/her answer were in the state:

$$|\psi\rangle_{asF} = |{\downarrow_z}\rangle_a |{\downarrow_z}\rangle_s |F\downarrow_z\rangle \tag{6}$$

The paradox lies exactly in the *discrepancy* between W's prediction about F's result, based on QM formalism (in particular on **(D)**), and the actual outcome obtained by F. From this thought experiment Wigner infers that, since QM *does not* apply to the human mind, the latter has to be the *cause* of the wave function collapse. Crucially, Wigner's friend paradox's does not hinge on a mismatch between the outcomes obtained by F's and W's measurements; rather, it is due to the incompatibility between F's result and W's *knowledge* of the system, grounded in QM formalism.

One of Wigner's students, Hugh Everett III, started from this paradox to develop his interpretation of QM (Barrett 2018). According to his approach, the collapse of the wave function never actually happens (i.e. it gets rid of the **Collapse** principle above): W's description of the quantum state of $a+s$, which involves a superposed state, is right. Still, F's claim that the spin of the system is $|{\downarrow_z}\rangle_s$ is not wrong since it is true in the branch in which F sits in. In other words, F describes a correct *relative* (to an observer) state of the system, whereas W's (equally right) prediction concerns the *univocal* state of the system before branching. In this interpretation, QM predictions are not violated, and the mismatch between W's knowledge and F's measurement is abridged insofar as they concern different perspectives, i.e. they diverge because they describe, in a certain sense, different systems.[7]

Crucially, one of the premises of Wigner's argument is the following:

(U) **Universality**: QM is a universal theory, that works for all microscopic and macroscopic systems as well.

Showing the failure of **(U)** seems to be the precise aim of Wigner's paradox, as explicitly acknowledged by Wigner (1961). Clearly, interpretations in the spirit of Everett do not need to reject **Universality**; indeed, they are usually presented as motivating the dismissal of the **Collapse** principle on acceptance of **(U)**.[8] Nonetheless, as we have just seen, these interpretations have an alternative way out of the paradox. At the same time though, all the interpretations

[7]As all the interpretations of QM, the Many Worlds approach comes in with different advantages and drawbacks. See: (Maudlin 2019, Ch. 6) for an accessible introduction to the interpretation, (Wallace 2012) for a modern statement of it, and (Barrett, Kent, Saunders & Wallace 2010) for a balanced discussion of its virtues and flaws.

[8]Deutsch (1985) came up with a thought experiment that should distinguish empirically in-

that reject **(U)** are not affected by Wigner's paradox (e.g. orthodox QM[9]): since they do not accept **Universality** in the first place, the argument does not take off from the very beginning. This is why Frauchiger and Renner do not assume **(U)** in their paradox, as they explicitly remark (2018, p.5). Indeed, since (von Neumann 1932), many have thought that something in the process of measurement must not be described quantum mechanically in order to perform the measurement itself successfully. The reason for that is, roughly, that unitary transformations[10] applied on a superposed state constrain the temporal evolution of the system to end in a new superposed state; this implies that a microscopical system that interacts with a measurement device (as big as one likes) will, if we apply only unitary evolutions, form a joint system in a superposition state (and so the measurement would not have a determinate outcome). Aware of this fact, the many founders of the theory already believed that QM could not be a universal theory.
As a matter of fact, Wigner's paradox does not favour only the Many Worlds interpretation of QM, but also the so-called 'QBism' (Fuchs 2010). According to Qbism, the wave function of a system corresponds to the piece of information that we have about the system itself. The quantum state of a system represents, according to QBists, the subjective probabilities that an epistemic agent would assign to its possible future experiences, i.e. the probabilities that he/she would assign to the possible outcomes of measuring the system. Albeit the probabilities assigned are subjective, they are not based on the agent's personal preferences. Indeed, they are grounded in an objective method that any rational agent must use,[11] that is to say, the formalism of QM. Still, according to QBism, nothing physically real corresponds to the wave function.[12] Rather,

terpretations that posit **(CP)** (like the just presented Wigner's perspective) and the Everettian's view. Interestingly enough, Frauchinger and Renner's paradox, as they (2018, p.2) explicitly acknowledge, has Deutsch (1985)'s and Hardy (1992, 1993)'s work as a starting point for their formulation.

[9]Frauchinger and Renner (2018, p.5) talk about 'most variants of the Copenhagen interpretation.' We prefer to use 'orthodox interpretation' for the formalism currently taught in physics courses in universities, and reserve the name 'Copenhagen interpretation' for the view endorsed by the founders of QM. Note that it is debatable if the founders of the theory endorsed a homogeneous view (Howard 2004).

[10]i.e. those standardly used in QM. A unitary evolution of a superposed state end with a new superposed state, because unitary operators preserve probabilities.

[11]For being considered *rational*; indeed, an agent interested in making accurate predictions ought to use a well-confirmed experimental theory, rather than personal preferences, in order to act rationally.

[12]Such an understanding of the wave function is not only opposed to the various forms

the wave function is understood as a guide for choosing reasonable subjective probabilities that have to be held about the outcome of possible measurements. As we said, QBism is favoured by Wigner's paradox exactly because it is not affected by the discrepancy between W's prediction and F's result: insofar as they have access to different knowledge, it goes without saying that the wave function that they attribute to the joint system diverges. Since our focus is mostly realist interpretations of QM, one might argue that QBism is outside the scope of the paper. Even though we think that there are reasons for considering QBism more an antirealist than a realist view of QM,[13] we acknowledge that some of its leading supporters strongly disagree. They argue that although the theory has an instrumental reading of the wave function, it is compatible with the idea that reality differs among different cognitive perspectives. Such a view, according to which there is an inter-dependence between epistemic agents and the world, has been called by some QBists 'participatory realism' (Fuchs 2017). We do not want to enter here into these controversial interpretational issues. Therefore, in what follows, we treat QBism as a realist view.

Finally, there is not much literature (that we know of at least) on how the Bohmian and GRW[14] interpretations of QM would deal with Wigner's paradox. Bohm's interpretation is not touched by the paradox, given the statistical character that measurements assume in the theory. Reasonably indeed, similar answers to those that have been put forward to Frauchinger and Renner's extended paradox (Lazarovici & Hubert 2019) can be rephrased toward the Wigner's original formulation of the thought experiment. Since the configuration of the particles upon which F and W perform a measurement is different, there is no contradiction in the fact that they obtain different results (i.e. Bohmian mechanics itself predicts that, since F and Wigner deal with different conditional wave functions, they obtain different results). In GRW, on the other hand, the evolution of a system is different from that of orthodox QM:

---

of wave function realism, according to which the wave function mathematically describes an existing field in the configuration space (Albert 1996, Ney 2013), a nomological entity (Goldstein & Zanghi 2013) or a distribution of properties (Monton 2004, Monton 2013, Dorato & Esfeld 2010), but also to other forms of instrumentalism (Bohr 1972–2006, Rovelli 1996).

[13]For example, realist interpretations of QM are united by the fact that they either accept that the wave function mathematically describes something existing in the mind-independent reality and (or) they posit a primitive ontology (Dürr, Goldstein & Zanghì 2012). As its supporters explicitly admit, QBism claims that the wave function has an epistemic, rather than an ontic, character; at the same time, it is not clear in what QBism's ontology consists in.

[14]GRW is not a single interpretation of QM but a family of models with different ontologies. In what follows, we will speak loosely about GRW interpretation without assuming any metaphysical picture in particular. For an introduction, see: (Ghirardi 2018).

a wave function may collapse spontaneously (i.e. even in the absence of any measurement). Moreover, the probability of a spontaneous collapse increases proportional to the number of particles that compose the system (and this is why, according to GRW, macroscopic entities, like tables and chairs, are never found in a superposed state). As a consequence, in GRW's view, according to Wigner the joint system of F and the silver atom is not in a superposed state before the friend's answer: given the dimension of such a system, the superposition would last just a tiny fraction of a second before collapsing spontaneously.
As we have seen, the main realist interpretations of QM have a way out of Wigner's paradox. Such a fact makes Wigner's thought experiment, as an argument in favour of the consciousness' role in the quantum collapse, extremely weak.[15]

3 Frauchinger and Renner's extended paradox

In contrast to the original Wigner's paradox, Frauchinger and Renner's version has the form of a 'no-go theorem': it assumes some principles, and it shows how the assumptions imply a contradiction. This in turns points to the rejection of one of the assumptions as the only solution of the paradox. For other notable theorems in QM, the principles assumed are so intuitive and reasonable that the choice of dropping one of them cannot be taken light-heartedly. Indeed, deciding which principle must be dropped has a very significant impact on how the world, depicted by the interpretation, is and how our knowledge of it works. The possibility of comparing how the different interpretations deal with the paradox is one of the most relevant features of Frauchinger and Renner's paradox. We postpone such an evaluation to (§4). In this section instead, we rewrite both the assumptions of the paradox (§3.1) and the contradiction itself (§3.2), with the aid of epistemic multi-agent logic (Fagin et al. 2003). Such a job will show not only how some assumptions are based on the theorems of these logics, but it will also pinpoint clearly where the contradiction arises. The choice of using this type of logic is motivated by the fact that Frauchinger and Renner's formulation of the assumption is epistemological. The basic symbols

[15] Indeed, as pointed out by Esfeld (1999) and Ballentine (2019), Wigner changes idea about the paradox, both for philosophical and physical reasons, starting from Zeh (1971)'s first attempts at finding a physical explanation of the wave function collapse.

that we will use are the standard primitive of epistemic logic, '$K_a p$',[16] which stands for 'agent a knows that p', the conjunction 'and' ('$\wedge$'), the negation 'it is not the case that' ('$\neg$'), and the material implication 'if x, then y' ('$x \rightarrow y$'). Finally, we will make modest use of the modal logic operators 'it is necessary that p' ('$\Box p$') and 'it is possible that p' ('$\Diamond p$'). As is usual, we assume that the operator 'K' is governed by axioms **T (Factivity)**:

(F) Factivity: $Kp \rightarrow p$.

and **4 (KK)**, which guarantees that the accessibility relation between possible worlds is reflexive and transitive. Finally, we assume that the modal and epistemic accessibility relations are different, and, following Hintikka (1962, Ch.1), that the two operators' interaction is governed by the following axioms: $Kp \rightarrow \Diamond p$ and $\Box p \rightarrow Kp$.

3.1 The extended paradox's assumptions

The extended formulation of the paradox takes its beginnings from three epistemic principles. The first assumption is the following:

> Suppose that agent A has established that:
>
> > *Statement* $\mathrm{A}^{(i)}$: "System s is in state $|\psi\rangle_s$ at time t_0."
>
> Suppose furthermore that agent A knows that:
>
> > *Statement* $\mathrm{A}^{(ii)}$: "The value x is obtained by a measurement of s w.r.t. the family $\{\pi_x^{t_0}\}_{x \in X}$ of Heisenberg operators relative to time t_0, which is completed at time t."
>
> If $\langle\Psi|\pi_\xi^{t_0}|\Psi\rangle = 1$ for some $\xi \in X$ then agent A can conclude that:
>
> > *Statement* $\mathrm{A}^{(iii)}$: "I am certain that $x = \xi$ at time t." (Frauchiger & Renner 2018, p.4)

Statement $\mathrm{A}^{(i)}$ simply tells us that the epistemic agent A prepared the system s in the state $|\psi\rangle_s$; it should be straightforward from that to infer:

$$K_A\,|\psi\rangle_s{}^{17} \tag{7}$$

[16]Note that it could be possible to further formalize Frauchinger and Renner's paradox with the common knowledge operator and its axioms; see: (van Benthem 2010, p. 140).

[17]Where, in this case, '$|\psi\rangle_s$' is a shortcoming for 'system s is in state $|\psi\rangle_s$'. In what follows we will be a bit sloppy writing directly, instead of a variable, the quantum state that the

In other words, since the agent has prepared the system in a determinate state, he/she knows that the system is in that particular state. The second part of the assumption is that x is the outcome of the measurement of the observable $\{\pi_x^{t_0}\}_{x \in X}$ performed at time t by A:

$$K_A(\pi_x^{t_0} = x \text{ at time } t)^{18} \tag{8}$$

In this case, it is safe to assume that this part of the assumption is uncontroversial: if the experiment performed personally by A gives as a result x, then it follows that A knows what the outcome is. Frauchinger and Renner explicitly assume that agent A knows how to apply the **Born rule** to any case. As a consequence, we can presuppose that A also knows how to deal with the particular case at hand:

$$K_A(\langle \Psi | \pi_\xi^{t_0} | \Psi \rangle = 1 \text{ for some } \xi \in X) \tag{9}$$

From (8) to (9) we can infer:

$$K_A(x = \xi \text{ at time } t) \tag{10}$$

In other words, the first assumption claims that if an epistemic agent knows how a system has been prepared, the results of an experiment carried by him/her and that **(BR)** predicts such an outcome, then he/she knows in what state the system is (according to QM). This first assumption, dubbed '**Q**' by the authors, can be summarized in a single formula:

(Q) **Quantum Reliability**: $[K_A \,|\psi\rangle_s \wedge K_A(\pi_x^{t_0} = x \text{ at time } t) \wedge K_A(\langle \Psi | \pi_\xi^{t_0} | \Psi \rangle = 1 \text{ for some } \xi \in X)] \rightarrow K_A(x = \xi \text{ at time } t)$

A way to simplify **(Q)** might be the following. If an agent A knows that the system s has been prepared in state $|\psi\rangle_s$, he/she knows how to apply **(BR)** and obtains, as a result of an experiment performed on s, the state $|x\rangle_s$; then A knows that QM predicts that s is in the state $|x\rangle_s$. Formally stated:

$$[K_A \,|\psi\rangle_s \wedge |Ax\rangle_s \wedge K_a(\mathbf{BR})] \rightarrow K_A \,|x\rangle_s \tag{\textbf{Q}}$$

epistemic agent knows. Albeit such a notation is imprecise, we thought that doing otherwise would have obliged us to introduce too many variables. To help the reader, we opted for this more intuitive, yet less precise, notation.

[18] For ease of exposition, we use '$\pi_x^{t_0} = x$' as a shortcut of 'x is the outcome of a measurement of $\pi_x^{t_0}$.'

Frauchinger and Renner assume that every agent involved in the paradox knows how to apply QM as well as other relevant information about the experiment (and therefore in which state the system has been prepared and how to apply **(BR)**). Therefore, $K_A |\psi\rangle_s$ and K_a**(BR)** will always hold for every epistemic agent involved. Therefore, we can omit them in our formulation of **(Q)** to streamline it a little. Therefore we obtain that **(Q)**, in the economy of the paradox, grants that when an epistemic agent A performs a measurement and obtains an outcome x on a system s ($|Ax\rangle_s$), he/she knows that the system is in the state corresponding to the results of its measurement ($K_A |x\rangle_s$); in our formalism again:

$$|Ax\rangle_s \rightarrow K_A |x\rangle_s \qquad (\mathbf{Q}^\dagger)$$

According to Frauchinger and Renner (2018, p.2), **(Q)** should 'captures the universal validity of quantum theory (or, more specifically, that an agent can be certain that a given proposition holds whenever the quantum-mechanical Born rule assigns probability-1 to it).' As already mentioned, **(Q)** diverges from **(U)** because it does not demand that an observer can describe himself/herself using quantum formalism.

The second explicit assumption of the paradox is the following:

> Suppose that agent A has established that:
>
> > *Statement* A$^{(i)}$: "I am certain that agent A', upon reasoning within the same theory as the one I am using, is certain that $x = \xi$ at time t."
>
> Then agent A can conclude that:
>
> > *Statement* A$^{(ii)}$: "I am certain that $x = \xi$ at time t." (Frauchiger & Renner 2018, p.5)

Rewritten, once again, with the aid of epistemic logic, it looks like: [19]

(C) **Outcome Transmissibility**: $K_A(K_B(MQ \wedge x = \xi \text{ at time } t)) \rightarrow K_A(x = \xi \text{ at time } t)$

In our simplified notation:

$$K_A(K_B(MQ \wedge |\xi\rangle_s)) \rightarrow K_A |\xi\rangle_s \qquad (\mathbf{C})$$

[19] Where 'MQ' stands for 'in what consists in and how to correctly use the formalism of QM.'

As is the case with **Q**, since Frauchinger and Renner assume that every agent involved in the paradox knows QM and how to make predictions with it (so for every epistemic agent A of the paradox, $K_A MQ$ holds), we can omit it in our formulation of **C**. Accordingly, we will use in our demonstration:

$$K_A(K_B\,|\xi\rangle_s) \rightarrow K_A\,|\xi\rangle_s \qquad (\mathbf{C}^\dagger)$$

It is easy to see that $(\mathbf{C}^\dagger)$ can be derived if we assume a well-known theorem of epistemic multi-agent logic:

$$K_A K_B p \rightarrow K_A p \qquad (\mathbf{K}_T)$$

Insofar as **(C)** is easily derivable from an essential theorem of epistemic multi-agent logic $(\mathbf{K}_T)$, that relies in turn on few assumptions (Hintikka 1962),[20] we must conclude that the former is a highly plausible epistemic principle.
We will now move to the last (explicit) assumption of Frauchinger and Renner's formulation; we quote them (2018, p. 5) extensively again:

> Suppose that agent A has established that:
>
> *Statement* A$^{(i)}$: "I am certain that $x = \xi$ at time t."
>
> Then agent A must necessarily deny that:
>
> *Statement* A$^{(ii)}$: "I am certain that $x \neq \xi$ at time t."

As with **(C)**, this assumption is a conditional. The antecedent (Statement A$^{(i)}$) seems to be:

$$K_A(x = \xi \text{ at time } t) \qquad (11)$$

where the consequent (Statement A$^{(ii)}$) is:

$$\neg K_A \neg (x = \xi \text{ at time } t) \qquad (12)$$

In our simplified notation:

$$K_A\,|\xi\rangle_s \rightarrow \neg K_A \neg\,|\xi\rangle_s \qquad (13)$$

The two authors state this conditional (**(S)** hereafter), as the epistemic prohibition that the very same observable might assume simultaneously different

[20] $(\mathbf{K}_T)$ can be easily demonstrated in any system that accepts **(F)**, assuming *ad absurdum* that '$K_A K_B p$' and '$K_A p$' hold.

values.[21] We can rewrite its logical form as:

(S) Outcome Uniqueness: $K_A p \rightarrow \neg K_A \neg p$

The logical expression '$\neg K_A \neg p$' means that among the worlds epistemically accessible by A, there is at least one in which p; as such, this formula is weaker than '$K_A p$'. Since the antecedent is logically stronger than the consequent, **(S)** results in a straightforward theorem in every epistemic logic. In the economy of the thought experiment, condition **(S)** imposes that, 'from the viewpoint of an agent who carries out a particular measurement, this measurement has one single outcome' (Frauchiger & Renner 2018, p. 2).[22]

Other than the three principles assumed by Frauchinger and Renner, their derivation relies on some other implicit assumptions. Renouncing one of them is a possible way out of the paradox, but we will not take into consideration such a move.[23] For the sake of completeness, we will report here these assumptions. The authors accept that every epistemic agent involved in the thought experiment (F, F^*, W, W^*, see §3.2 below) knows all the relevant information that concerns the experiment, i.e. they know how the whole apparatus is set up and of what the experiment consists. Moreover, they know the three principles just mentioned (**(Q)**, **(C)** and **(S)** and every other assumption of the paradox itself); they also know that other agents know these assumptions as well[24] and that they can make inferences through the *modus ponens*. Finally, two significant implicit assumptions are endorsed, as pointed out by Nurgalieva and del Rio (2019). First, that 'agent's *memories* are ultimately a physical system. In particular, they are quantum systems'

[21] It is worth noting that Renner stated principle **(S)** differently during a conference (The recordings can be found at: https://video.ethz.ch/speakers/its/2018/autumn/colloquium.html.): 'It is not possible that I am certain that the result of an experiment is z and not z at the same time.' It can be translated in the formalism of epistemic logic as: **($S^\dagger$)** $\neg \Diamond K_A (p \wedge \neg p)$. Replacing **S** with **($S^\dagger$)** does not have any bearing on their paradox since it would still be possible to derive the contradiction. **($S^\dagger$)** is much stronger than **(S)**, since rejecting it entails that the very notion of 'epistemic access' loses its very meaning. As such, the violation of **($S^\dagger$)** would make a system of epistemic logic trivial and full of contradictions. Since **(S)** can be deduced from **($S^\dagger$)**, we consider in what follows only the rejection of **(S)**.

[22] From a semantical point of view, **(S)** means that the frame of worlds is serial, that is that from each world there is at least one accessible.

[23] This does not imply that they are uncontroversial or that they cannot be resisted. Cf. Nurgalieva and del Rio (2019).

[24] So they are more likely idealized knowers rather than real human beings. Probably one can say that these are "common knowledges in a logical sense, that is that everyone knows that everyone knows *ad infinitum*, but this is not relevant here for us.

(Nurgalieva & del Rio 2019, p.273). Second, a weaker form of **(U)** is required, according to which an agent A may "model measurements performed by any other agent B as reversible evolutions in B's lab", that is "all agents are considering the evolution of another agents in their labs unitary" (Nurgalieva & del Rio 2019, p. 272).

3.2 Frauchinger and Renner's paradox

Having clarified the assumptions of the formulation of the thought experiment proposed by Frauchinger and Renner, we now turn to its presentation. Their version of the paradox involves four epistemic agents: instead of Wigner and his friend, the extended paradox involves two epistemic agents playing Wigner's role, W and W^*, and two of their friends, F and F^*, carrying out the experiment in two different laboratories (L and L^*). F^* makes the experiment in the laboratory L^*, with W^* waiting outside. At the same time, F and W are respectively inside and outside L (see fig. 2). Moreover, L and L^* are far from each other.

The thought experiment starts when F^* uses a quantum random generator, whose unique possible outcomes are $|heads\rangle$, with probability $\frac{1}{3}$, and $|tails\rangle$, with probability $\frac{2}{3}$. If F^* obtains $|heads\rangle$, then he/she prepares the system s in the state $|\downarrow\rangle_s$ while if he/she obtains $|tails\rangle$, he/she prepares it in a superposed state ($|superposed\rangle$ henceforth) between $|\uparrow\rangle_s$ plus $|\downarrow\rangle_s$ (each with a probability of 50%). Having prepared the system according to the random generator outcome, F^* sends the system s to F, who in turn measures it on the base $|\uparrow\rangle_s$ and $|\downarrow\rangle_s$; it is easy to see that F will obtain two times out of three $|\downarrow\rangle_s$ and $|\uparrow\rangle_s$ in the remaining third.

The following steps are slightly more technical. W^*, that is to say, the one outside F^*'s laboratory, measures L^*'s state on the following base:

$$|+\rangle_{L^*} = |F^*heads\rangle + |F^*tails\rangle \tag{14}$$

$$|-\rangle_{L^*} = |F^*heads\rangle - |F^*tails\rangle \tag{15}$$

Where $|F^*heads\rangle$ stands for the fact that F^* finds $|heads\rangle$ with his/her random generator and $|F^*tails\rangle$ stands for the fact that F^* finds $|tails\rangle$. The other Wigner, W, performs an analogous measurement on L on the following base:

$$|+\rangle_L = |F\uparrow\rangle_s + |F\downarrow\rangle_s \tag{16}$$

$$|-\rangle_L = |F\uparrow\rangle_s - |F\downarrow\rangle_s \tag{17}$$

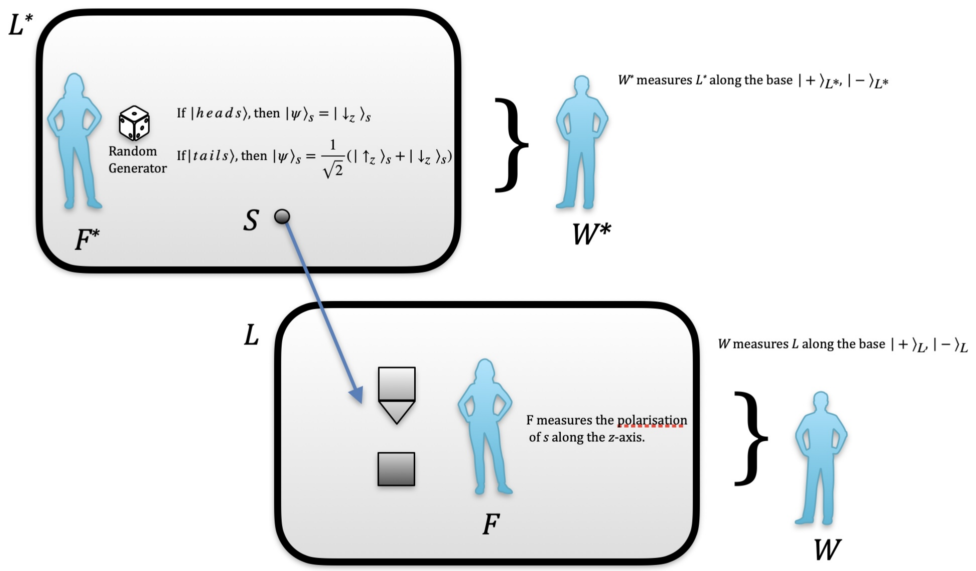


Figure 2: **Frauchinger and Renner's Paradox.** There are two agents, F and F^*, that make measurements inside two different laboratories, L and L^*. Two agents, W and $W*$ wait outside, respectively, L and L^*. Inside L^*, $F*$ prepares a system s in a state determined by a (quantum random) generator. The system s is then sent to L, where F performs a measurement on s. W and W^*, at different times, make measurements upon, respectively, composite systems L^*F^*s and LFs. By comparing the knowledge that every agent should have about the quantum systems involved, one can derive that the knowledge of one of the agents is incoherent.

Once again, $|F\downarrow\rangle_s$ (or $|F\uparrow\rangle_s$) means that $|\downarrow\rangle_s$ (or $|\uparrow\rangle_s$) is the outcome of the measurement performed by F on s. Applying **(Q)**, **(C)** and **(S)**, it can be easily shown in a few passages that the following four statements hold:

(a) F^* knows that, if $|tails\rangle$ is the outcome of his/her random generator, then W will find, performing a measurement, $|+\rangle_L$.

(b) F knows that, if he/she finds $|\uparrow\rangle_s$, then F^*'s random generator outcome has been $|tails\rangle$.

(c) W^* knows that, if he finds $|-\rangle_{L^*}$, then F measured $|\uparrow\rangle_s$.

(d) W knows that, if he finds $|+\rangle_L$, then there are cases in which W^* measured $|-\rangle_{L^*}$.

We can rewrite them using our formalism:

$$K_{F^*}(|F^*tails\rangle \rightarrow |W+\rangle_L) \tag{a†}$$

$$K_F(|F\uparrow\rangle_s \rightarrow |F^*tails\rangle \tag{b†}$$

$$K_{W^*}(|W^*-\rangle_L \rightarrow |F\uparrow\rangle_s \tag{c†}$$

$$K_W(|W-\rangle_L \rightarrow \Diamond|W^*-\rangle_{L^*} \tag{d†}$$

We will now limit ourselves to justifying (a), insofar as the justification of (b), (c) and (d) would lead us too far from the main concerns of the present paper. For what follows, it is enough to say that the proof of (b) is similar to that of (a), and that (c) and (d) rely on the principles **(Q)** and **(C)**.[25]
Let us justify (a). If F^* finds $|tails\rangle$, then applying **(Q)** he/she knows that the state of the quantum random generator is $|tails\rangle$. The experiment is built such that he/she knows that if he/she obtains $|tails\rangle$, then he/she must prepare the system s in the state $|superposed\rangle$, before sending it to F. Since F^* knows **MQ**, he/she knows that a measurement performed by W on L will deliver the state $|+\rangle_L$ as a result. Applying *modus ponens* two times, F^* can derive from $|tails\rangle$ that $|W+\rangle_L$ holds. Therefore, F^* knows that if $|tails\rangle$ is the outcome of his/her quantum random generator, then W will measure $|+\rangle_L$, concluding the justification of (a†). Note that it is possible, applying **Factivity**, **(Q)** and *modus ponens*, to simplify further statements from (a†) to (d†). Doing so helps to pinpoint the contradiction reached in the paradox. Again, we show just how

[25] For an extended justification we invite the interested reader to see not only the original paper but also: (Bub 2017, Fortin & Lombardi 2019).

to simplify (a[†]), since the deduction of the others is almost identical. We start by applying **Factivity** to (a[†]), obtaining:

$$|F^{*}tails\rangle \rightarrow |W+\rangle_{L} \qquad (a_k, \mathbf{F}) \tag{18}$$

Now we consider the antecedent and, applying **(Q)** we obtain, from:

$$|F^{*}tails\rangle \tag{19}$$

its simplified version:

$$K_{F^{*}}\,|tails\rangle \quad (19, \mathbf{Q}, modus\ ponens) \tag{20}$$

Analogously we obtain from

$$|W+\rangle_{L} \tag{21}$$

its epistemic formulation:

$$K_{W}\,|+\rangle_{L} \quad (21, \mathbf{Q}, modus\ ponens) \tag{22}$$

These passages are enough for deriving (a[†])'s more intuitive form:

$$K_{F^{*}}\,|tails\rangle \rightarrow K_{W}\,|+\rangle_{L} \quad (18, 20, 22, \mathbf{Q}, modus\ ponens) \qquad (\mathrm{a}_1^{\dagger})$$

The result obtained after repeating the procedure with all the other conditionals, is:

$$K_{F^{*}}\,|tails\rangle \rightarrow K_{W}\,|W+\rangle_{L} \qquad (\mathrm{a}_1^{\dagger})$$

$$K_{F}\,|\uparrow\rangle_{s} \rightarrow K_{F^{*}}\,|tails\rangle \qquad (\mathrm{b}_1^{\dagger})$$

$$K_{W^{*}}\,|-\rangle_{L} \rightarrow K_{F}\,|\uparrow\rangle_{s} \qquad (\mathrm{c}_1^{\dagger})$$

$$K_{W}\,|-\rangle_{L} \rightarrow \Diamond K_{W^{*}}\,|-\rangle_{L^{*}} \qquad (\mathrm{d}_1^{\dagger})$$

Such a simplification is helpful in the derivation[26] of the contradiction that lies at the heart of the paradox. To formulate it, we start assuming the antecedent

[26] In his blog, Scott Aaronson (https://www.scottaaronson.com/blog/?p=3975) challenged some passages of this argumentation. Aaronson's critique has been developed in more detail by Healey (2018). In particular, they deny the counterfactual determinateness of quantum results by assuming universality **(U)** (Healey 2018, p. 1577). As we remarked above (cf. §2), Frauchinger and Renner do not assume **(U)** in their paradox; moreover, whether **(U)** has to be accepted is controversial (see, e.g., Dalla Chiara, (1977)). To avoid entering in this controversy, we left aside Aaronson and Healey's reply in what follows.

of ($d_1^\dagger$), and by assumption, we pick a case in which the consequent holds, i.e. a case in which '$K_{W^*}\,|-\rangle_{L^*}$' holds. Now, if $K_{W^*}\,|-\rangle_{L^*}$ is true, from ($c_1^\dagger$) we can infer using *modus ponens* that $K_F\,|\uparrow\rangle_s$. We repeat the process using this time $K_F\,|\uparrow\rangle_s$ and ($b_1^\dagger$), obtaining $K_{F^*}\,|tails\rangle$. Finally, we use *modus ponens* the last time with $K_{F^*}\,|tails\rangle$ and ($a_1^\dagger$) to derive $K_W\,|+\rangle_L$. So we have derived from $K_W\,|-\rangle_L$ that $K_W\,|+\rangle_L$ holds. However, if $K_W\,|+\rangle_L$ holds, then we may derive a contradiction using **(S)** ($K_a p \rightarrow \neg K_a \neg p$). Indeed, it is a QM rule that if measuring an observable with only two eigenvalues, when the system is in one eigenstate (that corresponds to one of the eigenvalues), then it must not be in the other. In other terms, $|+\rangle_L$ and $|-\rangle_L$ are mutually exclusive states,[27] so that:

$$K_W\,|+\rangle_L \rightarrow K_W \neg\,|-\rangle_L \tag{23}$$

must hold (and also $K_W\,|-\rangle_L \rightarrow K_W \neg\,|+\rangle_L$). Since from $K_W\,|-\rangle_L$ we derived $K_W\,|+\rangle_L$, we can say that:

$$K_W\,|-\rangle_L \rightarrow K_W\,|+\rangle_L \tag{24}$$

So that from (23) and (24) and *transitivity of implication* we obtain:

$$K_W\,|-\rangle_L \rightarrow K_W \neg\,|-\rangle_L \qquad (23, 24, \textit{transitivity of implication}) \tag{25}$$

Using **(S)**, we can derive from (25):

$$K_W\,|-\rangle_L \rightarrow (K_W \neg\,|-\rangle_L \wedge \neg K_W \neg\,|-\rangle_L) \qquad (25, \textbf{(S)}) \tag{26}$$

Since it is possible that $K_W\,|-\rangle_L$, we reach a contradiction; in these cases indeed, through *modus ponens*, we can infer from $K_W\,|-\rangle_L$ and (26) that:

$$K_W \neg\,|-\rangle_L \wedge \neg K_W \neg\,|-\rangle_L \qquad (\text{assumption}, 26, \textit{modus ponens}) \tag{27}$$

The last formula, (27), is contradictory: in epistemic logic, $K_A p \wedge \neg K_A p$ is a straightforward contradiction. Semantically, such a formula indeed would mean that in every world to which A has epistemic access, p must hold and yet there is at least a possible world accessible by A in which $\neg p$ is true.

[27] So that '$K_W\,|+\rangle_L \vee K_W\,|-\rangle_L$' must hold. From this, we can trivially derive '$K_W\,|+\rangle_L \rightarrow K_W \neg\,|-\rangle_L$'.

4 The consequences of the paradox on QM interpretations

As we have seen, Frauchinger and Renner's formulation reach a real contradiction, where Wigner's original version was mostly an argument against **(U)**. Moreover, the new paradox has another advantage over the original: the contradiction of W's knowledge concerns outcomes of measurement that *in principle* could be performed upon a system and not just a hypothetical formulation of the evolution of a system. To be fair, we do reckon that this feature is weakened by the fact that the experiment is, according to the present technologies, only theoretically and not practically feasible. Indeed, the experiment would require a system (such as a quantum computer), that can perform reversible measurements (i.e. avoiding the effect of decoherence) on macroscopic objects such as an entire laboratory.[28] Still, *gedankenexperiments* are valuable insofar as they can point out logical and foundational contradictions within the theory; and logical inconsistencies are independent of the practical possibility of performing them experimentally. Therefore, since its purely theoretical character is not as limiting as one might think, the thought experiment cannot be easily dismissed just as a toy example of no interest.
In this last section, we turn our attention to how different interpretations of QM cope with the paradox. We will focus only on those interpretations that reject the three main assumptions of the paradox (**(Q)**, **(S)** and **(C)**), dedicating a subsection to the possible rejection of each principle. In (§4.5), we will outline some alternative way out of the paradox.

4.1 Dropping (Q)

We start by examining those interpretations that, according to Frauchinger and Renner, already violate the reliability of QM prediction **(Q)**. According to the two authors, those interpretations that postulate a modification of the standard collapse (i.e. like GRW) and Bohm's interpretation should not accept **(Q)** in the first place. For what concerns GRW, we think that it is not clear whether it violates **(Q)**. The two authors do not justify such a claim, limiting themselves to noticing that 'these deviate from the standard theory already on microscopic scales, although the effects of the deviation typically only become noticeable in larger systems' (Frauchiger & Renner 2018, p. 5). It

[28]Note cursorily that the same premise is shared in the mentioned thought experiment proposed by Deutsch (1985). See fn.9 above.

is true that GRW and orthodox QM may differ in the description of some microscopical state and that **(BR)** is rejected. For instance, if for the latter measuring a superposed state ends with a determinate result, according to the former there is a tiny, yet existent, probability that it does not, i.e. that the state of the system after the interaction with a measuring device remains superposed. Still, if we understand the Born rule as an operational principle that allows us to make predictions on the outcome of a possible experiment, it holds in GRW interpretations. Indeed, GRW matches, at a microscopic level, the predictions of orthodox QM, and at a macroscopic level those of classical physics; it is at a *mesoscopic* level, instead, that GRW diverges with both (Bassi, Lochan, Satin, Singh & Ulbricht 2013). If there were some grounds for doubting that GRW rejects **(Q)**, it would be unclear nonetheless how GRW supporters would avoid the contradiction of Frauchinger and Renner's paradox. A possible way out might be that of mimicking GRW's answer to the original Wigner's friend paradox we saw above (Cf. §2). According to GRW, the wavefunction of a system has a chance to collapse in a determinate state spontaneously, and such a probability is proportional to the dimension of the system. GRW's explanation of why human beings never witness ordinary objects in superposed states is that objects composed of so many particles have a probability proximate to the certainty of collapsing even before any perceptive stimulus can be processed. As we have seen, Frauchinger and Renner's thought experiment involves four human beings divided into two laboratories; and it is required that they can send a quantum system from one laboratory to the other without changing its state. One may insist that, according to GRW, there is a too small probability that W and W^* might obtain a superposed state on objects as big as those contemplated in the thought experiment (i.e. L and L^*): as soon as the quantum system interacts with the first laboratory, the probability of its state collapsing are proximate to certainty. That the quantum state survives both interactions with these macroscopic entities is for all practical purposes almost impossible. This consideration would effectively block the contradiction of the paradox in the middle of its derivation. Yet, one may argue that such a reply ignores the fact that there is still a tiny probability that the quantum state will not spontaneously collapse, leaving the possibility of deriving Frauchinger and Renner's paradox. Whether friends of GRW might endorse our suggestion or Frauchinger and Renner's suggestion of dropping **(Q)** is tied to how strong their realist commitment to the theory is.

An analogous consideration might be proposed for Bohmian mechanics. Some authors[29] have challenged Frauchinger and Renner's claim that Bohm's interpretation violates **(Q)**. According to them, Bohmian mechanics would indeed rather reject (a†) and (c†): after F^* obtains a particular result with his/her random generator (let us say $|tails\rangle$), in the Bohmian's view we cannot ignore the other possibility (i.e. $|heads\rangle$). In a realist interpretation without collapse, as with Bohmian mechanics, a state like $|+\rangle_L$, that can be the outcome of a measurement performed by W, also contains the other possibility (i.e. $|-\rangle_L$), albeit as an empty wave. When we consider the pilot wave, the deduction proposed by Frauchinger and Renner does not hold anymore. Even if orthodox QM and the Bohmian interpretations have the same experimental results, still the metaphysical picture described is different. The resulting knowledge of a physical system that one epistemic agent may have, then, differs according to the two interpretations.

4.2 Dropping (S)

Frauchinger and Renner seem to implicitly argue that the more natural way to avoid the contradiction remains the rejection of **(S)** (i.e. that if an agent knows p, then he/she does not know $\neg p$). Indeed, they (2018, p. 6) state that 'although intuitive, **(S)** is not implied by the bare mathematical formalism of QM.' This might mean that, for an interpretation of QM, the rejection of such a principle is not a big deal, because it is not part of the theory itself. Moreover, they (2018, p. 6) argue that 'among the theories that abandon the assumption there are the [...] "Many Worlds interpretations".'[30]

We would like to challenge both of these claims. Let us begin with the latter (i.e. that Many Worlds violates **(S)**). They argue that approaches inspired by the renowned paper by Everett (1957) get rid of **(S)** because 'any quantum measurement results in branching into different "worlds", in each of which one of the possible measurement outcomes occurs' (Frauchiger & Renner 2018, p. 6). In order to show why the Many Worlds interpretation (allegedly) violates **(S)**, let us take the simple example proposed at the beginning, concerning a silver atom ('s') in a superposition between its z-spin eigenvalues. If an observer F measures the spin along the z-axis on the system s, according to the Many Worlds view, there will be a world w_1 in which F measures $|\uparrow_z\rangle_s$ and a world w_2 in which the F's twin, $F^\dagger$, measures $|\downarrow_z\rangle_s$; F knows, applying

[29] See for instance: (Sudbery 2017, Lazarovici & Hubert 2019).

[30] We are talking about the Many Worlds interpretation as well as the relative state formulation and their variants; see for example: (DeWitt 1970, Deutsch 1985, Albert & Loewer 1988).

(Q), that s is in state $|\uparrow_z\rangle_s$. Moreover, F knows, according to Everettian interpretation and applying **(C)**, that there is a twin of him/her that obtained in another branch $|\downarrow_z\rangle_s$. This clearly violates[31] what **(S)** would demand to hold, i.e.:

$$K_A |\uparrow_z\rangle \rightarrow \neg K_A \neg |\uparrow_z\rangle \tag{28}$$

This might be a way to argue that Many Worlds violates **(S)**. But such an explanation can be challenged. How to do so really boils down to how the theory is formulated (because it depends whether the version of the interpretation also accepts **(Q)** and **(C)**).[32] Still, we think there is a reasonable way for Many Worlds supporters to avoid rejecting **(S)**. Indeed, Many Worlds supporters could insist that during the measurements the world branches, i.e. in one branch F measures $|\uparrow_z\rangle$ and in the other, his/her twin (i.e. $F^\dagger$) measures $|\downarrow_z\rangle$. However, as F and $F^\dagger$ are not the very same person, neither are the two silver atoms the same atom. Therefore, it is false that F knows that system s is both in the states $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$. Instead, what he/she can know is that the silver atom s in his/her branch is in state $|\uparrow_z\rangle$, while $s^\dagger$, i.e. the silver atom in the other world measured by $F^\dagger$, is in state $|\downarrow_z\rangle$. By accepting that the two silver atoms are numerically different, no violation of **(S)** could be derived.

This explains why contrary to what Frauchinger and Renner affirm, i.e. 'Many Worlds interpretations manifestly violate (S)' (2018, p. 6), it is not straightforward that Many Worlds rejects **(S)**. Even conceding that supporters of *diverging* Many Worlds[33] may refuse to drop **(S)** as suggested above, one could argue that it is not the case for supporters of *overlapping* Many Worlds[34]. Even if we think that even in *overlapping* Many Worlds, a strategy like the one above could be defended, there is also a possible way to avoid rejecting **(S)**: claiming that '$K_A p$' is not a good way of modelling an agent A's knowledge because, in this interpretation of QM, what is known by an agent has to be relativized to the world in which it obtains.[35] As Nurgalieva and del Rio (2019) argued,

[31] Insofar as spin up and spin down on an axis are mutually exclusive states, such that the system can be only (at a time) in one eigenstate or in the other; from that, it is trivial to derive that if the system is in one of the eigenstates, then it is not in the other.

[32] Therefore, rather than our general discussion, one should take one variant of Many Worlds at a time; such a work cannot be accomplished for reasons of length.

[33] i.e. interpretation of the theory according to which 'macroscopic objects and events are always worldbound, each being part of one Everett world only' (Wilson 2020).

[34] i.e. interpretation of the theory according to which 'macroscopic objects and events may be part of several different Everett worlds' (Wilson 2020).

[35] For a semantic which fully develop such a proposal, see, e.g., (Wallace 2005, Saunders & Wallace 2008).

such logic would avoid Frauchinger and Renner's paradox without the need of dropping any of the assumptions above.
We challenge now the second claim above held by Frauchinger and Renner: dropping **(S)** is, independently of the interpretation assumed, the easiest way out of the paradox. One may think that **(S)** is just an intuitive principle as good as the other. Why we disagree will be soon explained. **(S)** is a trivial consequence of all systems of epistemic logic. Indeed, **(S)** is a particular instantiation of the more general axiom D, which is employed in any model of whichever modal logic. As a matter of fact, **(D)** is equivalent to:

$$\neg K_A(p \wedge \neg p) \tag{29}$$

which, in other words, is the requirement that no world is isolated, i.e. the frame of worlds is serial. Now, seriality is requested by **(F)**, which is by many considered the hallmark of *knowledge*. Since dropping **(S)** entails dropping **(F)**, renouncing to the former entails renouncing to modelling knowledge. Actually, it might be argued that it is even worse than that: even in models of doxastic logic, an instantiation of **(D)** is requested. Hence, a model which drops the counterpart of **(D)** is unfit to describe even the beliefs of epistemic agents, let alone their knowledge. Finally, it is relevant for the paper to note that dropping **(S)** entails dropping **(C)**. Indeed, 'an agent A knows p' being true means that p is true to every world to which A has access. If the frame of worlds is not serial, it is easy to find a counterexample of **(C)**. Take an agent A in w_1 which has access to w_2 and to the isolated world w_3. Suppose further that p is true in w_3 and false in w_2, and that there is an agent B at w_3. From the fact that A knows that B knows p, we cannot conclude that A knows p, insofar as there is at least a world (w_2) accessible by A in which p is false.
As we have seen, there are compelling reasons for not considering the three assumptions of Frauchinger and Renner's paradox as really on par. We argued so far that dropping **(S)** should be avoided, also because it implies dropping also **(C)** and **(F)**. Exactly because of that, one may wonder how viable is dropping one of the latter principles instead of **(S)**. The next two sections describe, respectively, the consequences of dropping **(C)** and **(F)**.

4.3 Dropping (C)

According to the authors, their version of the paradox seems to commit the Copenhagen interpretation to the refusal of **(C)**. Even setting aside the problem of whether or not this interpretation has ever been a coherent view in the first place, we are sceptical of the truth of such a statement. It is true indeed that the founders of QM were already conscious of the failure of **Universality** (and therefore the need for a weaker principle, such as **(U)**). Nonetheless, we do not see why they would have discarded **(C)** rather than **(S)**. From their side, Frauchinger and Renner do not seem to provide any reasons for that either.[36] Still, given that it is impossible for Copenhagen supporters to reply to the paradox, we think that wondering what they would have thought is not so interesting.[37] Instead, it is more worthwhile to investigate how modern variants of orthodox QM would deal with the paradox. Here the discussion proposed by Frauchinger and Renner is far more convincing, and generally speaking, the situation is more evident than in the previous cases. Indeed, it seems that relational interpretations (such as those proposed by Rovelli (1996)[38] deny the absolute character of quantum states; such a fact seems to imply that these interpretations discard **(C)**.[39] Something similar happens in approaches to measurement based on decoherence (Omnès 1992) since, according to such approaches, the outcome of a measure depends on how the measurement apparatus disperses coherence in the environment. In other words, there are huge differences (according to these interpretations) between, let us say, the measurement performed by W on L, that gives the result $|-\rangle_L$, and that carried out by F^*, with the result $|tails\rangle$. Now, W infers the contradiction comparing the results he obtained with the one found by F^*. However, such a comparison is not allowed in these interpretations because, without using **(C)**, W should limit himself to the analysis of results obtained by his measuring apparatus. As the two authors note in the appendix, this fact is evident in the approaches to decoherence such as Consistent Histories. It is straightforward to show that the history typical of the thought experiment analyzed in the present paper,

(HS) F^* observes that $|tails\rangle$, F observes that $|\uparrow\rangle_s$, W^* observes that $|-\rangle_{L^*}$, W observes that $|-\rangle_L$

should have the probability of one out of twelve ($\frac{1}{12}$), where (HS)'s partial

[36]Arguably, the theoretical price of dropping **(S)** is much higher than that of **(C)**, so what the authors claim has a certain plausibility. But such a consideration still does not justify the claim that the supporters of the Copenhagen interpretation would have made such a move.

[37]Albeit there has been an interesting attempt of doing so in (Bub 2017).

[38](Laudisa & Rovelli 2019) is a good introduction to the relational interpretation of quantum mechanics; moreover, in section 2.5, the authors briefly argue that Frauchinger and Renner's paradox can be seen as an indirect support to this interpretation of quantum mechanics.

[39]To face this feature, Yang (2019, §3.1) proposed a modified version of the relational interpretation that explicitly avoids Frauchinger and Renner's paradox by accepting what he dubbed the "synchronization principle".

history:

(HS†) F^* observes that $|tails\rangle$, W observes that $|-\rangle_L$

should have a zero probability (0%). Such a result is impossible, insofar as (HS†) is a part of (HS). Indeed, this is not what happens in Consistent Histories approaches, since in these interpretations (HS) and (HS†) do not belong to the same framework; but if this is so, then in a Consistent Histories understanding of QM, the transmission of knowledge necessary fails. We think that this is enough for showing that this family of interpretations would not accept **(C)** as we have presented it.

The QBist interpretation deserves a separate treatment. In order to avoid the contradiction, QBism has to reject, as pointed out by the authors, principle **(C)**. Insofar as QBism is an epistemic approach to QM, we think that it has, at least, to tell a story about such a choice. This issue is quite pressing since the strongest point of QBism is precisely the dismissal of ontological headaches[40] through a substantial introduction of Bayesianism. One might argue that a full formalization of Bayesianism is a form of dynamic epistemic logic. If this is true, the employment of this type of logic is not for free, since it brings along with itself most of the structure of epistemic logic. Therefore, the choice of getting rid of **(C)** cannot be taken with a light heart.[41]

Note though that there are two 'easy' ways out of the impasse. Firstly, as pointed out by Nurgalieva and del Rio (2019, p. 275), QBism is explicitly a single-agent theory (i.e. it deals with the knowledge of a single agent at a time). Therefore, a QBist could insist that it is unfair to analyze the no-go theorem in the QBist interpretation before a generalization of the interpretation has been proposed. The second easy justification of QBists' dismissal of **(C)** might come from an explicit rejection of realism. As we have seen, one of the premises for the derivation of **(C)** was **Knowledge Transmissibility**; we, at least, see such a principle to be tightly connected with some kind of principle of reality: if Alice and Bob both know how an object x truly is, either: (i) there is a way in which $x$really is, and Alice's and Bob's knowledge of it is identical (so **($\mathbf{K}_T$)** must hold), or (ii) their knowledge diverges, but then (since both of their knowledge is true) there is no determinate mind-independent[42] way in which x is (or, if it exists, epistemic agents cannot have access to it). It is clear

[40] Such as the metaphysical status of the wave function, the spooky actions at a distance, collapse etc.

[41] Even though this is the official position advocated by the main defenders of QBism (DeBrota, Fuchs & Schack 2020).

[42] Or perspective-, language-, and so forth, independent.

that (ii) violates $(\mathbf{K}_T)$: insofar as if Alice's and Bob's knowledge diverge, from the fact that Bob knows that Alice knows that the object x is such and such, it does not follow that also Bob knows that the object x is such and such (rather, he will know that the object x is different from Alice's description).[43]

4.4 Dropping (F)

We have seen in (§4.2) that it is better to drop **(C)** rather than **(S)**, since rejecting the latter also requires renouncing the former. Still, dropping **(C)** is highly problematic itself, insofar as it can be done only by also dropping **Factivity**.[44] Allowing that **Factivity** does not hold in some cases is complicated for a wide variety of reasons. The first is that rejecting **(F)** implies that:

$$K_A p \wedge K_A \neg p \tag{29}$$

is not contradictory; indeed, formulae like (29) are a contradiction only in logical systems that, through **(F)**, can derive from it:

$$p \wedge \neg p \tag{30}$$

Since the contradiction reached in Frauchinger and Renner's paradox (27) is of the form of (29), if an interpretation of QM drops **(F)**, it avoids the contradiction. Therefore, dropping **Factivity** allows one to avoid Frauchinger and Renner's paradox without being committed to dropping (C). Instead of the latter, a QM interpretation should rather drop **(F)** directly. But what does dropping **(F)** mean?

If **(F)** is not an axiom then, from a semantic point of view, it is possible that a knower does not have epistemic access to his/her actual world. This is why many think that **(F)** is necessary for a good understanding of what knowledge is. It seems that the very concept of knowledge that if one *knows* that a given state of affairs obtains, then it really holds. In other words, necessary features for knowing that p are both the fact that one believes that p and that

[43] Another possibility might be following Frauchinger and Renner (2018, p. 9)'s suggestion and model their 'is certain that' with a certain degree of belief, and let it 'be replaced by something like "would arbitrarily large amount on".' This would concretely mean to model their assumption with a primitive different from the one we used ('$K_A p$'). Another possibility might be to replace **(C)** with a weaker principle; this is a route that, as Frauchinger and Renner point out, is currently under investigation by J.B. DeBrota, C.A. Fuchs, and R. Schack.

[44] It can be trivially shown that **(C)** could be derived from **(F)**. Moreover, we remind the reader that dropping **(C)** entails also rejecting **KK**.

p itself is true. Therefore, since knowledge deals with true claims, many would doubt that dropping **Factivity** would allow one to continue to talk about knowledge; rather, they would insist that dropping **(F)** means talking about belief or credence rather than knowledge itself.[45]
Yet in the quantum context, it might be that rejecting **Factivity** might make sense in antirealist interpretations. If QM does not describe the world, but is rather a computational device for making accurate predictions, say, then accepting that one knows that the system is in a certain state does not imply that the system really is in that state. Of course, this implies accepting a different[46] notion of truth, one that makes the fact that one knows something true even if this something does not really exist. In other words, extremely antirealist interpretations have a simple way out of Frauchinger and Renner's paradox in the form of dropping **Factivity** in the quantum mechanical context: it might be true that the outcome of an experiment is that, say, s is in $|\downarrow_z\rangle_s$ even if there is no property of having spin in the world, or even no s. The fact that an agent knows something is usually understood as the fact that the agent knows something of the world. **Factivity** delivers exactly such a connection between epistemic agents and reality. It goes without saying then that, in an instrumentalist interpretation of QM, according to which one must not believe that the theory describes the world itself, one can 'know' the QM description of a system without committing himself/herself to talking about the world. Rejecting **Factivity** in the context of QM delivers precisely the attitude of antirealist interpretations.

4.5 Logical ways out

Be that it as it may, there have also been suggested logical solutions to the paradox. Fortin and Lombardi (2019), for instance, propose avoiding the contradiction by rejecting classical logic. They claim that Frauchinger and Renner's argument is one of the many demonstrations that quantum objects violate Boolean logic and follow a non-distributive one. According to them, this new version of Wigner's paradox is not groundbreaking, insofar as it is just

[45]Note though that Boge (2019) shows that even in a system of doxastic logic, Frauchinger's and Renner's paradox is a theorem. Since in these logical system **(F)** is not an axiom, it is debatable whether rejecting it really helps avoiding Wigner's extended paradox. However, Boge (2019) proves the contradiction assuming a doxastic epistemic logic of kind **KD45**. It could be argued, following Hintikka (1962), that **5** cannot be an axiom of any reasonable doxastic logic.

[46]Different from the mainstream account of truth as correspondence.

the umpteenth proof that classical logic is of no use when we deal with QM. Still, we think that such a thesis raises, in this context, many doubts. To sum them up briefly, there is an enormous difference between belief and knowledge about a physical system and the system itself. Arguably, that physical system's behaviours violate classical logic might not be by itself problematic;[47] but that also our knowledge violates it is rather more perplexing. We do think that the rejection of non-distributivity, as Fortin and Lombardi present it, is highly controversial in the case of epistemic logic. Rather than being just another proof that quantum objects cannot be modelled with classical logic, Frauchinger and Renner's paradox raises a deep puzzle about the relationship between the logic that describes quantum objects and describes our knowledge of them.[48]

As a concluding remark, we would like to point out a final, and more promising way, to circumvent the paradox logically. The strategy consists of claiming that '$K_A p$' is not an accurate way to formalize knowledge in QM. According to some interpretations, indeed, it might be natural to claim that the knowledge of a quantum system is relative to a particular environmental context (it may be an observer, a world or a set of temporally indexed events, and so on). It is, for example, what we proposed early on for avoiding the fact that Many Worlds violates **(S)**, but it seems reasonable to use this strategy also for QBism and Consistent Histories approaches. This possibility consists of parameterizing the proposition known in respect to a particular frame of reference (logically cashed out as a set of possible worlds) and creating a new logical system with it. Intuitively, from a syntactic point of view, it could be possible to have a knowledge operator indexed not only from the point of view of the agents, but also with respect to the world in which they inhabit.[49] Another

[47]For instance, non-relativistic QM might violate classical logic for what concerns the identity of a quantum system. For an extensive overview, see: (French, Krause, Decio et al. 2006).

[48]For instance, if we assume **(F)** and that quantum objects are modelled by quantum logic, one can immediately see that epistemic logic cannot be classical. In quantum logic, distributivity does not hold: from $q(p \vee \neg p)$ one cannot deduce $(q \wedge p) \vee (q \wedge \neg p)$. However, distributivity holds in epistemic logic; by activity, an agent that knows $q(p \vee \neg p)$ could derive $(q \wedge p) \vee (q \wedge \neg p)$, thus contradicting quantum logic. Such a deep problem, which relies on many physical and metaphysical assumptions, will not be discussed here. We thank an anonymous referee for their interesting remarks on this point.

[49]We will not develop this idea further; moreover, it is difficult to find in the literature a ready-made logical system, at least that we know, that would be helpful in this context. Therefore we suspect that a new logical system should be developed in order to cash out how knowledge works for these interpretations of QM. In particular, for those versions of Many

promising way to circumvent the paradox might be to combine an indexical semantics (Kaplan 1989) with an epistemic logic: such a union should be able to cash out the idea that the truth value of a sentence like 'W knows that the system is in state $|\psi\rangle$' depends on the context of utterance (it might be a determinate observer in a particular region of space, or in a different world and so on). Although we think that the just sketched possibilities might supply a way out of Frauchinger and Renner's paradox without being committed to the dismissal of realism, no definite conclusion can yet be drawn. Indeed, insofar as such a logic (and its application to a quantum context) will not be developed in each particular interpretation of QM, there are no a priori reasons for thinking that it really would work. Nonetheless, there are good reasons for being optimistic. Nurgalieva and del Rio (2019, pp. 282-284) independently proposed a parametrized possible world semantic that, we think, should do the work we are proposing here. Indeed, they successfully showed that, in this kind of logic, Frauchinger and Renner's paradox might be easily avoided. The next step should be that of showing how the interpretations above could use this logic to model how epistemic agents have knowledge of quantum systems.

5 Conclusions

Throughout the paper, we have shown why Frauchinger and Renner's version of the paradox is far more cogent that Wigner's original formulation. As we have seen, Wigner's paradox is an argument in favour of the failure of **Universality** and, in the author's intention, of the role of consciousness in quantum collapse. Frauchinger and Renner's reformulation instead reaches a true contradiction assuming three highly intuitive and reasonable principles. Formalizing them in epistemic logic terms has been not only a natural choice but also a really fruitful one. Indeed, it has been helpful for clarifying logical connections, highlighting hidden premises and elucidating the sustaining structure of the paradox itself. Moreover, we think that our presentation of the paradox achieved two remarkable results: it is a better analysis of the costs associated with the rejection of the principles assumed and of the bearing that the paradox has on the main interpretations of QM.
As we have seen, **(Q)** amounts to the acceptance that QM can make accurate predictions using the **Born Rule**. Given the incredibly substantial experimen-

---

Worlds (that accept both **(Q)**, **(S)** and **(C)**), such a work is pressing insofar as rejecting realism (and **(F)**) does not seem to be a viable realist option.

tal results gathered in past years, we think that there is (almost) no way to reject this premise. At the same time, **(S)** is a consequence of most systems of epistemic logic whose rejection has catastrophic consequences on a logical system. Therefore, contrary to what Frauchinger and Renner's seemed to imply, dropping **(C)**, rather than **(S)**, is the less problematic choice. Still, the refusal of **(C)** cannot be accepted easily. Indeed, dropping **(C)** entails also the failure of **Factivity**, that is (almost) universally accepted as the hallmark of knowledge itself. We briefly presented how the rejection of **Factivity** might make sense in the context of QM. Indeed, we argued that **Factivity** could fail in strong antirealist interpretations of the theory. Since rejecting **(S)** entails dropping **(C)**, and the dismissal of the latter requires abandoning **(F)**, the prospects of being realist about QM seem dim: if it is so implausible to reject **(Q)**, and the rejection of **(C)** and **(S)** entails a dismissal of **Factivity**, one might think that realism must go. Indeed, it seems that it is the only way out. Given how intuitive the assumptions of the paradox are and the nefarious consequences that follow dropping them, we conclude that Frauchinger and Renner's paradox highly encourages an antirealist interpretation of QM. Both **Factivity** and **(C)** are, in fact, expression of some *principle of reality*, i.e. some principle connected with the idea of an objective and mind-independent reality to which different epistemic agents can have access.

That interpreting QM realistically raises many problems has been known since the first formulation of the theory. For example, the very fact that quantum systems cannot have definite values for all their properties at the same time, and the fact that such indeterminacy is a trivial consequence of the mathematical formalism adopted, cast suspicion on the idea that the theory completely describes the microscopic world. Such a tension between realism and QM is certainly not something new, but it is nowadays less mainstream than it was sixty years ago. The introduction of realist interpretations of QM and the 'triumph' of scientific realism over the dominant Neo-positivists view of science, have softened the idea that QM favours antirealism. Nonetheless, our conclusion is that the relevance of Frauchinger and Renner's paradox consists exactly in the fact that not only can it be seen as an argument that favours an antirealist over a realist view of QM, but it also commits, prima facie, realist interpretations to the rejection of strongly realist-oriented principles. However, even if the paradox is a challenge for realists, we are confident that some answers are available. As we have seen, the main interpretations of QM have a way out of the paradox.

Contrary to what argued by Frauchinger and Renner, our feeling is that it is controversial that GRW must drop (at least operationally) **(Q)**. Rather, GRW can avoid the paradox by rejecting the possibility of measuring the superposed state of huge systems, like those considered in the thought experiment (i.e. L and L^*). According to GRW, a superposed system that big would almost instantaneously collapse on one of its terms. At the same time, Frauchinger and Renner's claim that Bohmian mechanics rejects also **(Q)** has been challenged in the literature; their committal to the empty wave indeed blocks Frauchinger and Renner's paradox in the middle of its derivation. As a matter of fact, then, accepting some form of GRW or Bohmian mechanics is the more straightforward realist way of avoiding the paradox.

Still, we agree with them that interpretations such as QBism and Consistent Histories are naturally seen as renouncing principle **(C)**. Different from what Frauchinger and Renner claim, we think that the Many Worlds view is committed to the same choice. On the one hand, we have shown that it is not straightforward that Many Worlds and its variants explicitly violate **(S)**. On the other, insofar as Many Worlds' violation of **(S)** can be derived through **(C)**, it seems more reasonable to drop the latter rather than the former, especially given how crucial **(S)** is in most systems of epistemic logic. If we are right in claiming that renouncing these principles lead to an antirealist view, one may conclude that QBism, Consistent Histories and Many Worlds are doomed to antirealism. We argued at length against this conclusion. These interpretations indeed could insist that quantum knowledge must be indexed (either to epistemic agents or to branches), putting forward an alternative logical system. Such a work has been sketched by Nurgalieva and del Rio. It shows already that logical systems of this kind avoid Frauchinger and Renner's paradox. Even though only deeper investigations can provide certain conclusions, we think what has been shown is enough for a (cautious) optimism.

Since we think that the experimental success of QM in a wide variety of different contexts cries out for a realist interpretation of the theory, we claim that finding the answers above is as pressing as ever. Frauchinger and Renner's paradox, like many other no-go theorems and paradoxes in quantum foundations, has as the easiest way out embracing an antirealist view on QM. But as we have argued, the last word has not been said. All the main realist interpretations of QM have some resources for challenging the paradox.

Acknowledgements

We would like to thank Valia Allori, Matteo Morganti and Frida Trotter for helpful comments and discussions on previous drafts of the paper.

References

Albert, D. & Loewer, B. (1988), ‘Interpreting the many worlds interpretation’, *Synthese* **77**(2), 195–213.
URL: *http://www.jstor.org/stable/20116589*

Albert, D. Z. (1992), *Quantum Mechanics and Experience*, Harvard Up.

Albert, D. Z. (1996), Elementary quantum metaphysics, *in* J. T. Cushing, A. Fine & S. Goldstein, eds, ‘Bohmian Mechanics and Quantum Theory: An Appraisal’, Springer Netherlands, Dordrecht, pp. 277–284.

Ballentine, L. E. (2019), ‘A meeting with wigner’, *Foundations of Physics* **49**(8), 783–785.
URL: *https://doi.org/10.1007/s10701-019-00283-x*

Barrett, J. (2018), Everett’s relative-state formulation of quantum mechanics, *in* E. N. Zalta, ed., ‘The Stanford Encyclopedia of Philosophy’, winter 2018 edn, Metaphysics Research Lab, Stanford University.

Barrett, J., Kent, A., Saunders, S. & Wallace, D. (2010), *Many Worlds?: Everett, Quantum Theory, and Reality*, Oxford University Press.

Bassi, A., Lochan, K., Satin, S., Singh, T. P. & Ulbricht, H. (2013), ‘Models of wave-function collapse, underlying theories, and experimental tests’, *Reviews of Modern Physics* **85**(2), 471–527.
URL: *http://dx.doi.org/10.1103/RevModPhys.85.471*

Boge, F. J. (2019), ‘Quantum information versus epistemic logic: An analysis of the frauchiger–renner theorem’, *Foundations of Physics* **49**(10), 1143–1165.
URL: *http://dx.doi.org/10.1007/s10701-019-00298-4*

Bohr, N. (1972–2006), *Collected Works, Vol. 1-12*, Amsterdam: Elsevier.

Bub, J. (2017), 'Why bohr was (mostly) right'.
URL: *https://arxiv.org/abs/1711.01604*

Chiara, M. L. D. (1977), 'Logical self reference, set theoretical paradoxes and the measurement problem in quantum mechanics', *Journal of Philosophical Logic* **6**(1), 331–347.

DeBrota, J. B., Fuchs, C. A. & Schack, R. (2020), 'Respecting one's fellow: Qbism's analysis of wigner's friend', *Foundations of Physics* **50**(12), 1859–1874.

Deutsch, D. (1985), 'Quantum theory as a universal physical theory', *International Journal of Theoretical Physics* **24**(1), 1–41.
URL: *https://doi.org/10.1007/BF00670071*

DeWitt, B. S. (1970), 'Quantum mechanics and reality', *Physics today* **23**(9), 30–35.

Dorato, M. & Esfeld, M. (2010), 'Grw as an ontology of dispositions', *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **41**(1), 41–49.

Dürr, D., Goldstein, S. & Zanghì, N. (2012), *Quantum physics without quantum philosophy*, Springer Science & Business Media.

Esfeld, M. (1999), 'Essay review wigner's view of physical reality michael esfeld1', *Studies in History and Philosophy of Modern Physics* **30**, 145–154.

Everett, H. (1957), '"relative state" formulation of quantum mechanics', *Rev. Mod. Phys.* **29**, 454–462.
URL: *https://link.aps.org/doi/10.1103/RevModPhys.29.454*

Fagin, R., Moses, Y., Halpern, J. Y. & Vardi, M. Y. (2003), *Reasoning about knowledge*, MIT press.

Fortin, S. & Lombardi, O. (2019), 'Wigner and his many friends: A new no-go result?'.
URL: *https://arxiv.org/abs/1904.07412*

Frauchiger, D. & Renner, R. (2018), 'Quantum theory cannot consistently describe the use of itself', *Nature Communications* **9**(1), 3711–3721.
URL: *https://doi.org/10.1038/s41467-018-05739-8*

French, S., Krause, D., Decio, K. et al. (2006), *Identity in physics: A historical, philosophical, and formal analysis*, Oxford University Press.

Fuchs, C. A. (2010), 'Qbism, the perimeter of quantum bayesianism'.
URL: *https://arxiv.org/abs/1003.5209*

Fuchs, C. A. (2017), On participatory realism, *in* I. T. Durham & D. Rickles, eds, 'Information and Interaction: Eddington, Wheeler, and the Limits of Knowledge', Springer International Publishing, Cham, pp. 113–134.

Ghirardi, G. (2018), Collapse theories, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', fall 2018 edn, Metaphysics Research Lab, Stanford University.

Goldstein, S. & Zanghi, N. (2013), 'Reality and the role of the wave function in quantum theory', *The wave function: Essays on the metaphysics of quantum mechanics* pp. 91–109.

Hardy, L. (1992), 'Quantum mechanics, local realistic theories, and lorentz-invariant realistic theories', *Phys. Rev. Lett.* **68**, 2981–2984.
URL: *https://link.aps.org/doi/10.1103/PhysRevLett.68.2981*

Hardy, L. (1993), 'Nonlocality for two particles without inequalities for almost all entangled states', *Phys. Rev. Lett.* **71**, 1665–1668.
URL: *https://link.aps.org/doi/10.1103/PhysRevLett.71.1665*

Healey, R. (2018), 'Quantum theory and the limits of objectivity', *Foundations of Physics* **48**(11), 1568–1589.

Hintikka, J. (1962), *Knowledge and Belief*, Ithaca: Cornell University Press.

Howard, D. (2004), 'Who invented the "copenhagen interpretation"? a study in mythology', *Philosophy of Science* **71**(5), 669–682.

Kaplan, D. (1989), Demonstratives: An essay on the semantics, logic, metaphysics and epistemology of demonstratives and other indexicals, *in* J. Almog, J. Perry & H. Wettstein, eds, 'Themes From Kaplan', Oxford University Press, pp. 481–563.

Laudisa, F. & Rovelli, C. (2019), Relational quantum mechanics, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', winter 2019 edn, Metaphysics Research Lab, Stanford University.

Lazarovici, D. & Hubert, M. (2019), 'How quantum mechanics can consistently describe the use of itself', *Scientific Reports* **470**(9), 1–8.

Maudlin, T. (2019), *Philosophy of Physics: Quantum Theory*, Princeton Foundations of Contemporary Philosophy, Princeton University Press.
URL: *https://books.google.it/books?id=uY6ADwAAQBAJ*

Monton, B. (2004), Quantum mechanics and 3n-dimensional space. wave function ontology, wavefunction realism.
URL: *http://philsci-archive.pitt.edu/2090/*

Monton, B. (2013), Against 3n-dimensional space, *in* D. A. A. Ney, ed., 'The Wave Function: Essays in the Metaphysics of Quantum Mechanics', Oxford University Press.

Ney, A. (2013), Ontological reduction and the wave function ontology, *in* A. Ney & D. Z. Albert, eds, 'The Wave Function: Essays on the Metaphysics of Quantum Mechanics', Oxford University Press, pp. 168–183.

Nurgalieva, N. & del Rio, L. (2019), 'Inadequacy of modal logic in quantum settings', *Electronic Proceedings in Theoretical Computer Science* **287**, 267–297.
URL: *http://dx.doi.org/10.4204/EPTCS.287.16*

Omnès, R. (1992), 'Consistent interpretations of quantum mechanics', *Rev. Mod. Phys.* **64**, 339–382.
URL: *https://link.aps.org/doi/10.1103/RevModPhys.64.339*

Rovelli, C. (1996), 'Relational quantum mechanics', *International Journal of Theoretical Physics* **35**(8), 1637–1678.

Saunders, S. & Wallace, D. (2008), 'Branching and uncertainty', *The British Journal for the Philosophy of Science* **59**(3), 293–305.

Sudbery, A. (2017), 'Single-world theory of the extended wigner's friend experiment', *Foundations of Physics* **47**(5), 658–669.
URL: *https://doi.org/10.1007/s10701-017-0082-7*

van Benthem, J. (2010), 'Modal logic for open minds', *CSLI lecture notes* (199).

von Neumann, J. (1932), *Mathematische Grundlagen Der Quantenmechanik*, J. Springer.

Wallace, D. (2005), 'Language use in a branching universe'.
URL: *http://philsci-archive.pitt.edu/2554/*

Wallace, D. (2012), *The emergent multiverse: Quantum theory according to the Everett interpretation*, Oxford University Press.

Wigner, E. P. (1961), Remarks on the mind-body question, *in* I. J. Good, ed., 'The Scientist Speculates', Heineman, pp. 284–302.

Wilson, A. (2020), *The Nature of Contingency: Quantum Physics as Modal Realism*, Oxford, UK: Oxford University Press.

Yang, J. M. (2019), 'Consistent descriptions of quantum measurement', *Foundations of Physics* **49**(11), 1306–1324.
URL: *https://doi.org/10.1007/s10701-019-00305-8*

Zeh, H. (1971), On the irreversibility of time and observation in quantum theory, *in* B. d'Espagnat, ed., 'Foundations of quantum mechanics', Academic Press New York, pp. 263–273.