Can we Rule out the Need for Consciousness in Quantum Mechanics?

J. Acacio de Barros[a], Gary Oas[b]

[a]*School of Humanities and Liberal Studies*
San Francisco State University
San Francisco, CA 94132
[b]*Stanford Pre-Collegiate Studies*
Stanford University
Stanford, CA 94305

Abstract

In this paper we examine some proposals to disprove the hypothesis that the interaction between mind and matter causes the collapse of the wave function, showing that such proposals are fundamentally flawed. We then describe a general experimental setup retaining the key features of the ones examined, and show that even a more general case is inadequate to disprove the mind-matter collapse hypothesis. Finally, we use our setup provided to argue that, under some reasonable assumptions about consciousness, such hypothesis is unfalsifiable.

Keywords: Measurement problem; von Neumann-Wigner interpretation; collapse of the wave function; fourth-order interference

1. Introduction

One of the central issues within Quantum Mechanics (QM) is the measurement problem. Though many different solutions to it have been offered (e.g. [1–6]), there is no consensus among physicists that a satisfactory resolution has been achieved. Perhaps the main reason for this disagreement is the lack of clear experimental procedures that could distinguish an interpretation from another; in fact. For example, Bohm's theory yields exactly the same predictions as the standard Copenhagen interpretation for quantum systems [7], at least for most measurable quantum systems[1].

Among the proposed solutions, perhaps one of the most controversial is von Neumann's idea that a measurement is the result of the interaction of a (conscious) mind with matter [11]. This idea posits two distinct types of dynamics for quantum systems: one linear, to which all matter is subject under its standard evolution, and another non-linear and probabilistic, to which matter is subject when it interacts with the observer's

Email addresses: `barros@sfsu.edu` (J. Acacio de Barros), `oas@stanford.edu` (Gary Oas)

[1]For extreme cases where there might be some differences, albeit not necessarily directly observable; see [8, 9] or [10].

mind. This is a substance-dualist view, where matter and mind exist in different realms and satisfy different laws of Nature. This interpretation has Henry Stapp as its currently best-known supporter [12]. We shall also call the hypothesis that the interaction with a mind causes the collapse of the wave function the Consciousness Causes Collapse Hypothesis (CCCH).

Recently, some authors claimed that the CCCH was inconsistent with already available empirical evidence (see, e.g. [13, 14]). In this paper, we examine the CCCH with respect to such claims, in particular those of [13], and show that their proposal does not provide a way to falsify the CCCH. We then modify their proposal to a stripped-down version that retains the main features of an experiment needed to falsify the CCCH. This exposes a fundamental problem: to test the CCCH one would need to make a conscious being part of the experimental setup. Unless we subscribe to a panpsychist view of consciousness (which the CCCH proponents usually do not), such types of experiment pose a fundamental problem: to have a conscious being, one needs reasonably high temperatures (compared to absolute zero). Thus, any experiment that distinguishes two orthogonal states of a measurement, as we shall see is necessary, cannot be brought to its original quantum state, as this would imply controlling all the quantum states in a thermal bath. Therefore, For All Practical Purposes (FAPP), the outcomes of such experiments would be inconclusive, and they would not test the CCCH. In fact, this suggests that, due to environmental decoherence, the CCCH is unfalsifiable FAPP.

We organize this paper in the following way. In Section 2 we briefly discuss the von Neumann interpretation of quantum mechanics. In Section 3 we present Yu and Nikolic's experiment, and describe why it does not work as proposed. Then, in Section 4, we modify their experimental setup, and analyze under which conditions the modified experiment needs to be performed to test the CCCH. We end the paper with some conclusions.

2. The von Neumann collapse interpretation of QM

In this section, we present the idea behind von Neumann's interpretation. To do so, we start with a quick statement of the famous measurement problem, which motivates his collapse theory. In his seminal book [15], von Neumann starts with the assumption that every physical system can be represented as a vector $|\psi\rangle$ in a Hilbert space $\mathcal{H}$. This representation is one-to-one, in the sense that not only every system has a corresponding vector, but that to every vector there is, in principle, a corresponding system. Observable quantities are represented in this Hilbert space as linear Hermitian operators. The spectral decomposition theorem tells us that a Hermitian operator $\hat{A}$ can be written as

$$\hat{A} = \sum_i a_i \hat{P}_i,$$

where $a_i \in \mathbb{R}$ and $\hat{P}_i$ are projection operators such that $\hat{P}_i \hat{P}_j = \delta_{ij} \hat{P}_j$. In von Neumann's view, the dynamics of a system is more complicated, and we should distinguish two types. One type is given when the system does not interact with a measurement device. When this is the case, the evolution of the state $|\psi\rangle$ follows a deterministic and linear evolution given by Schroedinger's equation. Namely, the state of the system at time $t_1 \geq t_0$ is given by

$$|\psi(t_1)\rangle = \hat{U}(t_1; t_0) |\psi(t_0)\rangle,$$

where $\hat{U}(t_1; t_0)$ is a unitary evolution operator between t_0 and t_1 given by

$$\hat{U}(t_1; t_0) = \exp\left[-\frac{i}{\hbar}\hat{H}(t_1 - t_0)\right],$$

and $\hat{H}$ is the Hamiltonian operator. If, on the other hand, the system interacts with a measurement device, the evolution is not linear nor determinist. During a measurement, each observable value a_i has a probability $p(a_i) = \left|\hat{P}_i|\psi\rangle\right|^2$ of being observed, and if the result of a measurement (with probability $p(a_i)$) is a_i, then wave-function collapses into a new state

$$|\psi\rangle \xrightarrow{a_i} \frac{\hat{P}_i|\psi\rangle}{\langle\psi|\hat{P}_i|\psi\rangle}.$$

So, according to this formulation, QM has two different types of evolution, one deterministic and one probabilistic; the former happens when there is no interaction with a measurement device, and the latter when such interaction occurs.

A natural question to ask within this theory is "what is a measurement device?" In principle, such a device, made out of "conventional" matter itself, should be describable by QM. Following von Neumann, let us assume this is the case, and let us have a Hilbert space $\mathcal{H} = \mathcal{H}_M \otimes \mathcal{H}_S$, where $\mathcal{H}_M$ is the space of the measurement device and $\mathcal{H}_S$ the space of the system being measured. Since we are considering this an isolated system, there is no interaction with an external measuring device (the device is part of the system itself). For simplicity, let us limit our measuring device to the observable

$$\hat{O} = \hat{P} - \left(\hat{1} - \hat{P}\right) = 2\hat{P} - \hat{1},$$

where $\hat{P}^2 = \hat{P} \neq \hat{1}$ is a projector, and $\hat{1}$ the identity operator. Clearly, $\hat{O}$ can have only two possible outcomes, $+1$ and -1. So, a measuring device for $\hat{O}$ needs to have the following properties. First, it should have a neutral state, its initial state, prepared to receive a system to be measured. We denote the neutral state of the measuring device by the vector $|\text{neutral}\rangle \in \mathcal{H}_M$. Second, the interaction of M and S should be such that the following evolution happens:

$$|\text{neutral}\rangle \otimes |+\rangle \to \hat{U}_{\text{int}}|\text{points to } +\rangle \otimes |+\rangle,$$

$$|\text{neutral}\rangle \otimes |-\rangle \to \hat{U}_{\text{int}}|\text{points to } -\rangle \otimes |-\rangle.$$

Here we represent the two possible final values of the measurement apparatus as either giving a measurement of "+" or "−," depending on the initial state of the system.

Since, according to QM, any linear superposition of states $|\pm\rangle \in \mathcal{H}_S$ are possible, what happens when we use the above interaction to measure superpositions? If we have the superposition

$$|\psi\rangle = c_+|+\rangle + c_-|-\rangle,$$

because $\hat{U}_{\text{int}}$ is linear, it follows that

$$|\text{neutral}\rangle \otimes |\psi\rangle \to c_+|\text{points to } +\rangle \otimes |+\rangle + c_-|\text{points to } -\rangle \otimes |-\rangle.$$

This seems to be exactly what we wanted: we end up with a correlation between $|\pm\rangle$ and the pointer's state $|\text{points to } \pm\rangle$. However, it is straightforward to see that the final state is *not* an eigenstate of either projector $\hat{1} \otimes |+\rangle\langle+|$ or $\hat{1} \otimes |-\rangle\langle-|$, and therefore does not correspond to an actual measurement, where an actual collapse happens. This contains the essence of the measurement problem: a system interacting with a measurement apparatus evolves according to a non-linear dynamics that is not the same for quantum systems.

So, von Neumann argued further, pointing out that if we started with a superposition, the interaction of S with a measurement apparatus M would result in a superposition. But we could think of another apparatus M' that measures M and S, and we'd still have a superposition, and keeping doing this indefinitely, ever adding more measurement apparatuses. We can even consider our eyes as a photodetector that measures this chain of measurement apparatuses, and we have no reason to assume, according to von Neumann, that we would not have a superposition. We can keep on going, including not only our eyes, but our optical nerves, up until we get to the brain, and we are left with a brain/measurement apparatus/system that is still in a superposition. That is intriguing, and since we never actually observe a superposition, this chain needs to stop somewhere. According to von Neumann there is only one step when we know *for sure* that we do not have a superposition: when we gain conscious knowledge of the measurement apparatus, i.e. when matter interacts with the mind. That is because we are never aware of observing any quantum superposition. He then proposed that the interaction between mind and matter causes matter to evolve probabilistically, according to Born's rule, and non-linearly. In other words, the mind causes the collapse of the wave function.

The CCCH is substance dualist. As it is well-known, dualist views of the mind suffer the problem of causal closure: how can the mind influence matter and *vice versa*? Though not directly addressing this issue, the CCCH states that the mind causes matter to behave differently, following a dynamics that is not the same when no interaction happens with a mind. So, in a certain sense, the CCCH postulates their interaction, albeit in a very specific way. The question remains as to whether this interaction may be used to actually provide a way for the mind to affect matter in a (consciously) controlled way.

Henry Stapp proposed a clever solution to this problem by using the "inverse" Quantum Zeno Effect (QZE) [16]. In the original QZE [17], it was shown that if we were to continuously observe an unstable particle, this particle would not decay. However, this argument can be modified, and it can be shown that by continuous and variable observations it is possible to force a particle to change a quantum state. So, in Stapp's example [16], if we start with a coherent state with amplitude α, *and* if our mind chooses to observe it, we end with a new amplitude $\beta > \alpha$, whereas if it chooses not to observe, the state maintains amplitude α.

Stapp [16] applies this idea to motor cortex measurements performed by Rubino, Robbins, and Hatsopoulos [18]. His idea is that, in the same way that the mind causes the collapse of the wave function, the effect of the mind "observing" a system can make it change its state from $|\alpha\rangle$ to $|\alpha + N\Delta\rangle$. There might be some (surmountable) problems with this model, discussed in more detail in [19, 20], but they are not relevant to the discussions here. What is important is to keep in mind that von Neumann's theory, though not popular among physicists and presenting some difficult philosophical challenges, not only solves the measurement problem, but also provides a possible way for the mind to affect matter, a major problem for substance dualists.

3. Proposed falsification of the CCCH

Since the CCCH attempts to be a scientific theory, it is reasonable to ask whether it is true or false. By true or false we of course mean whether there is supporting experimental evidence for it or if it can be or has been falsified, as we cannot, in a strict sense, prove a theory to be true. So, an interesting question is how can we try to falsify the CCCH.

In a recent paper [13], Yu and Nikolic argue that the CCCH has already been falsified, and proposed further modifications of a given experimental setup to make such conclusions beyond any reasonable doubt. Their argument starts with the idea that

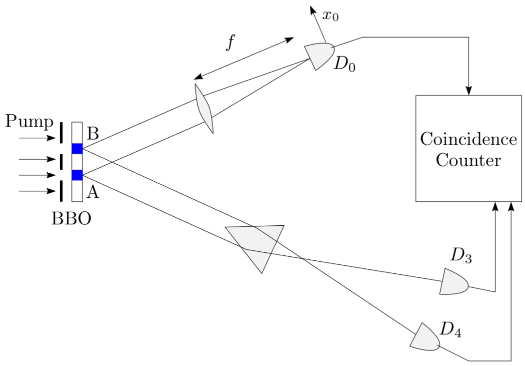
$$\text{CCCH} \rightarrow (\text{CWF} \iff \text{PR}),$$

where CWF is short for "collapse of the wave function" and PR for "phenomenal representation," i.e. the presence of phenomenal consciousness. Therefore, they argue, it if is possible to "observe" CWF without PR, the CCCH is falsified.

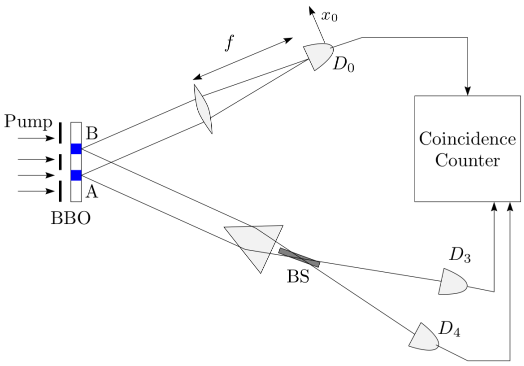
To understand Yu and Nikolic's argument, and our criticism of it, we need to look into the details of how they try to argue for the possibility of observing CWF without PR. They do so by using Kim et al.'s delayed choice experiment [21], which we now describe. In Kim et al., a pump laser beam impinges on a standard double slit, behind which a non-linear BBO crystal is placed. Through parametric down conversion, a pair of photons is generated in either region A or B of the crystal, residing behind each slit, with the signal photon going to a detector D_0 that can be moved to observe an interference pattern. The idler photon goes to either detectors D_3 or D_4 (Figure 1 (a)), thus allowing which-path information, or are scrambled in a beam splitter BS (Figure 1 (b)), erasing any which-path information, and again being detected on D_3 or D_4. Kim et al.'s experimental setup is shown in Figure 1.

To understand Kim et al.'s experiment, it is important to notice first that it is a fourth-order interference experiment[2]. Let us analyze what happens in each of the setups (for details relevant to the experiment discussed here, see, e.g. [24]). First, for the which-path information setup in Figure 1 (a), there is nothing strange. The pair of photons is produced either in A or B, and if it is produced in A the idler photon is detected in D_3, and if in B it is detected in D_4. Since the signal photon is produced in either A or B, the final probability of observing it in the variable-position detector D_0 is the same as the sum of the two probabilities, and shows no interference effect, as expected. For the interference setup shown in Figure 1 (b), things are more subtle, and the experimental setup resembles, conceptually, what happens with ghost interference (another fourth-order interference experiment) [25]. When the idler photons from A or B are joined, we loose which path information, but, more importantly, the idler side of the apparatus becomes an interference device itself, sensitive to momentum of the quantum state impinging on it. Different momenta, which are correlated with D_0, produce different interference patterns in D_0, and the overall probability distribution observed in D_0 is

[2]Readers not familiar with fourth-order interference are encouraged to consult [22] or one of the many excellent textbooks on quantum optics, such as [23].



(a)



(b)

Figure 1: Kim et al. experimental setup for the delayed choice quantum eraser [21].

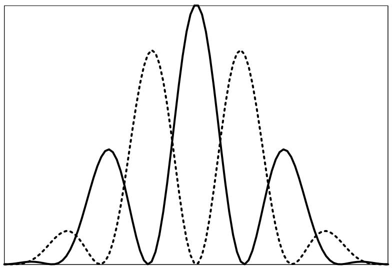


Figure 2: Probabilities $P\left(D_0, x|D_i\right)$ of observing a photon in detector D_0 positioned at x, conditioned on a detection on D_3 (solid line) or D_4 (dashed line).

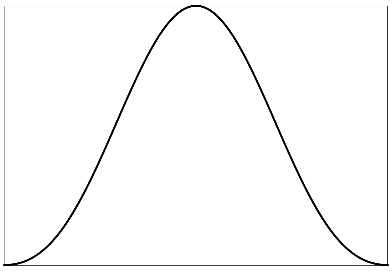


Figure 3: Probability as a function of x of observing a photon in detector D_0 positioned at x.

exactly the same as with setup (a). As a consequence, the *conditional* probability of detection on D_0 depends on a detection on D_3 or D_4 in the following way [21]:

$$P\left(D_0, x|D_3\right) = N\left(\alpha x\right)^{-2}\sin^2\left(\alpha x\right)\cos^2\left(\beta x\right), \tag{1}$$

$$P\left(D_0, x|D_4\right) = N\left(\alpha x\right)^{-2}\sin^2\left(\alpha x\right)\sin^2\left(\beta x\right), \tag{2}$$

where N is a normalization factor, and α and β parameters that depend on the optical geometry of the experiment and the correlated photons wavelength. The two conditional probabilities in (1) and (2) are shown in Figure 2. As we can see, by *conditionalizing* the data on the detection of, say, D_3, we observe an interference pattern, and likewise for the conditionalized data on D_4. However, as we can also see from Figure 2, the interference pattern obtained by conditionalizing on D_3 is shifted with respect to the one from D_4 (this is also clear from (1) and (2)).

That the interference pattern appearing in D_0 is conditioned to either D_3 or D_4 and is offset by $\pi/2$ for them is a crucial point: the interference pattern does not appear on D_0 without correlating it with the detections on D_3 or D_4. In fact, if we look at D_0, what we see is the unconditional $P\left(D_0, x\right)$, given by $P\left(D_0, x|D_3\right)P\left(D_3\right) + P\left(D_0, x|D_4\right)P\left(D_4\right)$, shown in Figure 3. If this were not the case, we would violate the no-signaling condition in quantum mechanics, as we could use a choice of detection apparatus in D_i to communicate instantaneously (or to the past) between an experimenter controlling D_i and another observing D_0. But since the observations are conditional, no violation of no-signaling occurs.

Returning to Yu and Nikolic's idea, their proposal was to use the human eye as a photodetector instead of D_i. This would not be an impossible task, given that human eyes are sensitive to single photons. As such, they argue that, in the which-path setup where the idler photon goes to D_3 and D_4, we could replace detectors D_3 and D_4 with a person observing the photons. If such observer were unconscious, then no collapse of the wave function would happen, and we would have an interference pattern on D_0. However, as we saw above, this proposal has a major difficulty: we would not get an interference pattern on D_0 with no collapse. As we saw above, to obtain an interference pattern, any which-path information about the idler photon needs not only to be erased by recombining the beams into an interferometer, but once recombined we would need to detect such photon and use coincidence counts to obtain the interference. And if we used an actual person to observe D_3 or D_4, such coincidence counts could only happen if such person was aware of the detection in their eye, as this would be required for knowing which detections in D_0 need to be counted. In other words, a human (or any other animal) used in this experimental setup would have to be aware of the detection of a photon within a certain window of time and be able to remember (or record on a piece of paper) such detection, such that later on an interference pattern could be reconstructed by coincidence counts[3].

But an interesting question is raised from Yu and Nikolic's proposal: could we falsify the CCCH with some device of this type? As we saw, their claim that the CCCH was (perhaps) already falsified is not correct, as their reliance on the quantum eraser experiment did not take into consideration the need for correlated counts. But perhaps some other version of the experiment could to it. In the next section we will show a general type of experiment to test the CCCH, and use it to argue that it is impossible to falsify the CCCH.

4. Is the CCCH falsifiable?

In this section we describe a proposed experiment to test the CCCH. The experiment we propose here is a natural extension of an earlier paper of Suppes & de Barros [26], and has the main features necessary to test the CCCH. Our goal here is not to propose a thought experiment, but to examine the characteristics of a realizable experiment, and discuss its conceptual and technical difficulties.

Since we want to test the CCCH, like Yu and Nikolic, we start with the eyes as photo-detectors. Nature provides us with exceptionally good photo-detectors in the kingdom of *Animalia* (see references in [26]). Of particular interest, is the fact that some insects have not only very efficient eyes (their efficiency is estimated to be between 40% and 78%), but very low dark-count rates (the locust *Schitocerca gregaria*, for example, has a dark-count rate of few photons per hour).

Perhaps one of the best candidates for such conditioning experiments is the cockroach (*Periplaneta americana*), for the following reasons [27]: it responds well to external stimuli for conditioning, it is well adapted to respond to very low-light environments (i.e. has

[3]In fact, the total number of photons reaching the participant (either human or not) is quite large, and it is not until coincidence counts are performed that this number is reduced. So, the task of reconstructing an interference pattern, even if the actual photon count per second could be reduced to a reasonable number to be dealt with, would be very time consuming and daunting.

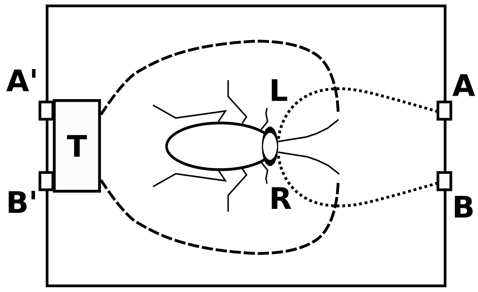


Figure 4: Proposed experimental setup. A photon impinges on A or B, and an optical fiber, represented by the dotted line, takes it to either the left (L) or right (R) eye, respectively. If a photon reaches L, the cockroach is conditioned to push a button at the end of a circuit (dashed line), and if the L button is pushed, a single photon is emitted at a precise and very short window of time.

good photo-detectors), and its neural circuitry is significantly easier to study compared to other well-known insects (such as the ubiquitous fruit fly). So, for that reason, in combination with the existence of successful conditioning experiments with insects, Suppes & de Barros [26] proposed that cockroaches could be classically conditioned to respond to single photons.

Here we assume that cockroach single-photon conditioning experiments could be successfully carried out, though probably there exists many technical difficulties (insect are not as easy to condition as some mammals). For our purpose, we will also assume that the cockroach is a conscious being. This is, of course, a controversial assumption, but the alternative would be to do our proposed experiment with more complex animals (say, humans). However, as it will become clear below, this assumption will not invalidate our conclusions, as they will apply to any animal.

The idealized experiment we propose is simple, and does not rely in entangled states (as does Kim et al.'s). Imagine we have a cockroach who has been conditioned to respond to single photons in the following way. If a photon impinges on the left eye of the cockroach, it moves its left antenna, whereas if a photon impinges on the right eye it moves its right antenna. The cockroach is then placed in a well isolated box where a photon can be sent to either the left or the right eye via optical fibers. If the cockroach's left antenna moves, the cockroach sends a signal to a device T that will generate a single photon from A'; if the right antenna moves, a single photon is generated from B'. Now, the idea here is that if instead of a single photon in A or B, a quantum superposition $|\psi\rangle = c_1|1\rangle_A|0\rangle_B + c_2|0\rangle_A|1\rangle_B$ was sent to the box, the output would be a quantum superposition if the cockroach is not conscious, whereas it would be a proper mixture if the cockroach caused a collapse of the wave function.

Now, to understand the experimental conditions necessary for such experiment to work, let us examine it in detail. We start with the Hilbert space of this setup, given by $\mathcal{H} = H_p \otimes H_c \otimes H_b \otimes H_{p'}$, where H_p is the Hilbert space for the impinging photon, H_c the cockroach, H_b the box itself (with all necessary devices), and $H_{p'}$ the outgoing photon. For example, when a single photon impinges on A, with

$$\rho_{1,0} = |1_A, 0_B\rangle\langle 1_A, 0_B|,$$

the initial state of the system is given by

$$\rho_{1,0} \otimes \rho^{\text{roach}}_{\text{ready}} \otimes \rho^{\text{box}}_{\text{ready}} \otimes \rho'_{0,0},$$

where

$$\rho^{\text{roach}}_{\text{ready}} = |\text{cockroach ready}\rangle\langle\text{cockroach ready}|,$$

$$\rho^{\text{box}}_{\text{ready}} = |\text{box ready}\rangle\langle\text{box ready}|,$$

and

$$\rho'_{0,0} = |0_{A'}, 0_{B'}\rangle\langle 0_{A'}, 0_{B'}|.$$

This system would evolve the following way:

$$\begin{aligned}
\rho_{1,0} \otimes \rho^{\text{roach}}_{\text{ready}} \otimes \rho^{\text{box}}_{\text{ready}} \otimes \rho'_{0,0} \to \\
\rho_{0,0} \otimes \rho^{\text{roach}}_{\text{left antennae}} \otimes \rho^{\text{box}}_{\text{ready}} \otimes \rho'_{0,0} \to \\
\rho_{0,0} \otimes \rho^{\text{roach}}_{\text{ready}} \otimes \rho^{\text{box}}_{\text{gen.photon A'}} \otimes \rho'_{0,0} \to \\
\rho_{0,0} \otimes \rho^{\text{roach}}_{\text{ready}} \otimes \rho^{\text{box}}_{\text{ready}} \otimes \rho'_{1,0},
\end{aligned}$$

where the label for the states should make them evident. A similar evolution would happen to $\rho_{0,1}$, leading to

$$\begin{aligned}
\rho_{0,1} \otimes \rho^{\text{roach}}_{\text{ready}} \otimes \rho^{\text{box}}_{\text{ready}} \otimes \rho'_{0,0} \to \\
\rho_{0,0} \otimes \rho^{\text{roach}}_{\text{ready}} \otimes \rho^{\text{box}}_{\text{ready}} \otimes \rho'_{0,1}.
\end{aligned}$$

Finally, if we started with a superposition given by, say, the state

$$\rho_{1,1} = \frac{1}{2}\left(|0_A, 1_B\rangle\langle 0_A, 1_B| + |1_A, 0_B\rangle\langle 0_A, 1_B| + |0_A, 1_B\rangle\langle 1_A, 0_B| + |1_A, 0_B\rangle\langle 1_A, 0_B|\right),$$

we would end with the linear evolution

$$\begin{aligned}
\rho_{1,1} \otimes \rho^{\text{roach}}_{\text{ready}} \otimes \rho^{\text{box}}_{\text{ready}} \otimes \rho'_{0,0} \to \\
\rho_{0,0} \otimes \rho^{\text{roach}}_{\text{ready}} \otimes \rho^{\text{box}}_{\text{ready}} \otimes \rho'_{1,1}.
\end{aligned}$$

Clearly, if the experiment could be performed like above, if the input is a superposition, we can take the partial trace over all other variables, and the output will also be a superposition. In other words, because the evolution is linear, the partial trace over $H_p \otimes H_c \otimes H_b$ of $\rho_{0,0} \otimes \rho^{\text{roach}}_{\text{ready}} \otimes \rho^{\text{box}}_{\text{ready}} \otimes \rho'_{1,1}$ would result in $\rho'_{1,1} \in H_{p'}$. However, if the cockroach's mind causes a collapse of the wave function inside the box, then the dynamics would not be linear, and the output would be the proper mixture

$$\rho'_{\text{mixture}} = \frac{1}{2}\left(|1_{A'}, 0_{B'}\rangle\langle 1_{A'}, 0_{B'}| + |0_{A'}, 1_{B'}\rangle\langle 0_{A'}, 1_{B'}|\right),$$

and not the pure state $\rho'_{1,1}$.

However, from the system's evolution above, we can see a major difficulty with such an experiment, which also will plague any other experiment attempting to falsify the CCCH. In order for a superposition to be detected at the output, the cockroach and box need to go back to its original *quantum* state. It is easy to see, for instance, that if the cockroach does not go back to its original state $\rho^{\text{roach}}_{\text{ready}}$, then the final state would be an entanglement between the different cockroach positions for inputs A or B. Then, if the outside experimenter observes this system (causing its collapse?), what they would see is a proper mixture, and not a superposition. Therefore, for such an experiment to work in testing the CCCH, the whole cockroach+box needs to be brought back to its original state. This means that every single atom that makes up the cockroach, for example, needs to be brought back to its original state. Of course, though a tremendously difficult task, it is not forbidden by quantum mechanics (though, not probably experimentally feasible, FAPP).

However, there is a deeper problem with this experiment. In order to perform such types of experiment with reasonable candidates for having phenomenal representation (a cockroach is already somewhat a questionable one), we need to include in the Hilbert space a thermal bath. The reason is that any candidate for phenomenal consciousness (unless we take a pansychist view) is a living creature who cannot have consciousness at temperatures close to absolute zero. Therefore, if we include the thermal bath on the description of the system above, even if we could bring the cockroach+apparatus back to its original quantum state, the outcome of the experiment would be irreversibly entangled with the thermal bath, and we would always observe at the end a proper mixture, *regardless* of whether the cockroach caused a collapse or not. Since a thermal bath is a necessary condition for a living candidate to have phenomenal consciousness, the CCCH is unfalsifiable.

5. Conclusions

The CCCH is arguably one of the most controversial solutions for the measurement problem in quantum mechanics, and it certainly does not share wide support within the foundations of physics community [28]. What perhaps makes it unappealing to most physicists is its substance-dualistic nature, where the existence of a mind that is not itself composed of ordinary physical matter gives the impression of a non-scientific theory.

The authors of this paper are not proponents of the CCCH, but it is important to recognize its main achievements. For instance, it does solve the measurement problem. Furthermore, it provides a substance-dualistic view of the mind that is not plagued by the causal closure problem that dates all the way back to Decartes. If such proposed closure mechanisms are correct, it also makes predictions about specific features of, say, the human brain, that are, after all, empirically observable. This, we believe, is a fascinating perspective, though we remain skeptical of its probability of success.

Be that as it may, it is not surprising that such a theory is often criticized, but mostly on metaphysical grounds (as are many of the different interpretations of QM). However, as we pointed out before, some researchers argued that the CCCH is not only metaphysically troubling, but also that it is empirically inadequate. This is, upfront, a strange claim, as von Neumann's interpretation gives exactly the same predictions as other interpretations of QM: i.e. they are empirically indistinguishable. However, if the mind plays a special rule on the measurement process, perhaps we can use this to create experiments where one could try to falsify the CCCH.

In this paper we examined one proposed experiment to disprove the CCCH proposed by Yu and Nicolic [13]. We saw that their proposal had a fatal flaw, as it did not consider the fact that to observe fourth-order interference requires coincidence counting. We then used this experiment as a springboard to a more general general framework for how to attempt to falsify the CCCH: produce an experimental setup where the non-linear nature of the quantum dynamics in the presence of consciousness can be distinguished from the linear dynamics in the absence of consciousness.

Another argument put forth against the CCCH was give by Thaheld [14]. In his paper, he used the Stark-Einstein law to argue that classical information is passed to the eye-brain system via absorption of photons by the retinal molecules. We will not go into the details of Thaheld's argument, since they are not required here, but we want only to point out that the classical information is passed because of an entanglement between a photon and the "classical" environmental variables, and also that the Stark-Einstein law assumes, deep down, a collapse of the wave function (either photon is absorbed by the molecule or not). Thaheld's argument against the CCCH also suffers from the same issues as the proposal put forth in Section 4.

Finally, we emphasize that, as we argued, any candidate for phenomenal consciousness, at least some consensus candidates, would have to be kept at their habitat's temperature. This implies that any such experiment would not be able to distinguish the linear from the non-linear dynamics, as we would always have an irreversible entanglement with a thermal bath. Therefore, any experiment trying to falsify the CCCH on the basis of its different dynamics seems doomed, FAPP.

Acknowledgments. This is a continuation of our work with Pat Suppes, who passed away in November 2014. This research was partially supported by the Patrick Suppes Gift Fund for the Suppes Brain Lab. Pat's support to this paper is gratefully acknowledged, as well as John Perry's hospitality while JAB visited CSLI at Stanford University, where part of this work was conducted.

References

[1] D. Bohm, A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables. I, Phys. Rev. 85 (2) (1952) 166–179. doi:10.1103/PhysRev.85.166.

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

[3] H. Everett, "Relative State" Formulation of Quantum Mechanics, Reviews of Modern Physics 29 (3) (1957) 454–462. doi:10.1103/RevModPhys.29.454.

[4] D. Bohm, David, J. Bub, A proposed solution of the measurement problem in quantum mechanics by a hidden variable theory, Reviews of Modern Physics 38 (3) (1966) 453–469. doi:10.1103/RevModPhys.38.453.

[5] R. Omnes, The Interpretation of Quantum Mechanics, Princeton University Press, 1994.

[6] C. A. Fuchs, Introducing QBism - Springer, in: M. Galavotti, D. Dieks, J. Gonzales, S. Hartmann, T. Uebel, M. Weber (Eds.), The Philosophy of Science in a European Perspective, New Directions in the Philosophy of Science, 2014.

[7] P. R. Holland, The quantum theory of motion: an account of the de Broglie-Bohm causal interpretation of quantum mechanics, Cambridge Univ Pr, 1995.

[8] J. A. de Barros, N. Pinto-Neto, The causal interpretation of quantum mechanics and the singularity problem and time issue in quantum cosmology, International Journal of Modern Physics D 7 (02) (1998) 201–213.

[9] J. A. de Barros, N. Pinto-Neto, M. A. Sagioro-Leal, The causal interpretation of dust and radiation fluid non-singular quantum cosmologies, Physics Letters A 241 (4) (1998) 229–239.

[10] A. Valentini, H. Westman, Dynamical origin of quantum probabilities, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 461 (2053) (2005) 253–272. doi:10.1098/rspa.2004.1394.

[11] J. von Neumann, Mathematical Foundations of Quantum Mechanics, Princeton University Press, Princeton, NJ, 1996.

[12] H. P. Stapp, Mind, Matter, and Quantum Mechanics, in: H. P. Stapp (Ed.), Mind, Matter and Quantum Mechanics, The Frontiers Collection, Springer Berlin Heidelberg, 2009, pp. 81–118.

[13] S. Yu, D. Nikolic, Quantum mechanics needs no consciousness, Annalen der Physik 523 (11) (2011) 931–938. doi:10.1002/andp.201100078.

[14] F. H. Thaheld, Can the Stark-Einstein law resolve the measurement problem from an animate perspective?, Biosystems 135 (2015) 50–54. doi:10.1016/j.biosystems.2015.07.005.

[15] J. v. Neumann, Mathematical Foundations of Quantum Mechanics, Princeton University Press, Princeton, NJ, 1932.

[16] H. P. Stapp, Mind, brain, and neuroscience, Cosmos and History 10 (1) (2014) 227–231.

[17] B. Misra, E. C. G. Sudarshan, The Zeno's paradox in quantum theory, Journal of Mathematical Physics 18 (4) (2008) 756–763. doi:10.1063/1.523304.

[18] D. Rubino, K. A. Robbins, N. G. Hatsopoulos, Propagating waves mediate information transfer in the motor cortex, Nature Neuroscience 9 (12) (2006) 1549–1557. doi:10.1038/nn1802.

[19] J. A. de Barros, G. Oas, Quantum mechanics & the brain, and some of its consequences, Cosmos and History: The Journal of Natural and Social Philosophy 11 (2) (2015) 146–153.

[20] J. A. de Barros, On a Model of Quantum Mechanics and the Mind, in: S. O'Nuallain (Ed.), Foundations of Mind, Cambridge Scholars Publishing, Berkeley, CA, 2016.

[21] Y.-H. Kim, R. Yu, S. P. Kulik, Y. Shih, M. O. Scully, Delayed "Choice" Quantum Eraser, Physical Review Letters 84 (1) (2000) 1–5. doi:10.1103/PhysRevLett.84.1.

[22] Z. Y. Ou, Quantum theory of fourth-order interference, Physical Review A 37 (5) (1988) 1607–1619. doi:10.1103/PhysRevA.37.1607.

[23] L. Mandel, E. Wolf, Optical Coherence and Quantum Optics, Cambridge University Press, 1995.

[24] M. H. Rubin, D. N. Klyshko, Y. H. Shih, A. V. Sergienko, Theory of two-photon entanglement in type-II optical parametric down-conversion, Physical Review A 50 (6) (1994) 5122–5133. doi:10.1103/PhysRevA.50.5122.

[25] D. V. Strekalov, A. V. Sergienko, D. N. Klyshko, Y. H. Shih, Observation of Two-Photon "Ghost" Interference and Diffraction, Physical Review Letters 74 (18) (1995) 3600–3603. doi:10.1103/PhysRevLett.74.3600.

[26] P. Suppes, J. A. de Barros, Quantum Mechanics and the Brain, in: Quantum Interaction: Papers from the AAAI Spring Symposium, Technical Report SS-07-08, AAAI Press, Menlo Park, CA, 2007, pp. 75–82.

[27] D. D. Lent, H.-W. Kwon, Antennal movements reveal associative learning in the American cockroach Periplaneta americana, Journal of Experimental Biology 207 (2) (2004) 369–375. doi:10.1242/jeb.00736.

[28] M. Schlosshauer, J. Kofler, A. Zeilinger, A snapshot of foundational attitudes toward quantum mechanics, Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 44 (3) (2013) 222–230. doi:10.1016/j.shpsb.2013.04.004.