

# The Nature of Belief in No-Collapse Everett Interpretations

*Paul Skokowski*

## 1. Introduction

In Chapter 4 of *The Quantum Mechanics of Minds and Worlds*, Jeffrey Barrett considers *the bare theory*, which takes the standard formulation of quantum mechanics (what he calls the von Neumann-Dirac formulation), and strips it of the collapse postulate. This yields an Everettian no-collapse interpretation of quantum mechanics in which “the usual deterministic linear dynamics always correctly describes the time-evolution of the quantum-mechanical state” (Barrett 1999, p. 94). There is much to agree with in Barrett’s discussion of the bare theory, and I concur with most every one of his conclusions about its shortcomings in the final sections as he closes out the chapter. The bare theory, as Barrett correctly points out, is at its most puzzling when it is applied to human mental states—in particular, human beliefs including introspective beliefs and experiences. And it is here where I think Barrett’s excellent analysis falls just a little short, not by any misapplication of quantum mechanical principles, but by not properly accounting for the complex nature of *human* belief states, and by this I mean in particular taking into account the neuroscience of belief (paying attention to its neural vehicles and causal roles), and paying due diligence to the contents and intentional properties of belief. By carefully accounting for these properties of belief within the quantum mechanical analysis, we will come to question two claims arising out of the bare theory about an observer of a superposition. First, when asked in a certain way about the result of the experiment, that the observer actually believes what he reports, and second that the observer would end up with a disjunctive belief about the experiment. In both cases we shall see that no beliefs actually have the content reported, and so the agent has no such beliefs. And the reasons for this have to do with paying close

attention to the vehicles, contents, and causal roles of the beliefs in question. We will find that this is a side-benefit to taking into account details of the neuroscience and content of belief.

## 2. The Bare Theory, Introspection, and Superposition

As an example of what one is confronted with in the bare theory, Barrett considers an observer  $M$  of a Stern Gerlach measurement that will measure the  $x$ -spin of a spin  $1/2$  system  $S$ .<sup>1</sup> Before interacting with the device, the system  $S$  is in an eigenstate of  $z$ -spin, and the observer  $M$  is in an eigenstate of being ready to measure the  $x$ -spin of the system  $S$ . After the measurement, the composite system of the observer  $M$  and the spin  $1/2$  system  $S$ , according to the bare theory, will be given by Barrett's equation as follows:

$$|\psi\rangle = 1/\sqrt{2}(|x\text{-spin up}\rangle_M |\uparrow_x\rangle_S + |x\text{-spin down}\rangle_M |\downarrow_x\rangle_S).$$

Or, more compactly,

$$|\psi\rangle = 1/\sqrt{2}(|\uparrow\rangle_M |\uparrow_x\rangle_S + |\downarrow\rangle_M |\downarrow_x\rangle_S).$$

Since we are talking about belief and experience, it will be important to recognize that the observer  $M$ , who is a competent observer of experiments, is a human observer. Barrett does initially develop a model of an automaton that represents spin measurements, where "this model requires a close correspondence between physical memory configurations and mental states..." (Barrett 1999, p. 95). But ultimately, it is the human experience of a superposition that needs explaining. For example, in the first sentence of the section entitled *The Account of Experience*, Barrett asks, "So just how far can the bare theory go in explaining our experience?" (Barrett 1999, p. 110). We will concentrate then on the nature of human belief when it is involved in measuring a superposition.

Now let us consider Barrett's original example of an observer of the outcome of an experiment that results in a superposition. Barrett proposes asking the observer, who is in the superpositional state, "Did you get a determinate result of either  $x$ -spin up or  $x$ -spin down?" (Barrett 1999, p. 98).

<sup>1</sup> We can consider the particles to be electrons for the purposes of this paper.

Note that when a person is asked to report about the content of a belief they hold, they need to introspect in order to access that belief and so be able to report the content of that belief. Barrett agrees, saying that *M* “would believe that he knows what the result is” (Barrett 1999, p. 98). Such a *belief* of *M*’s about what he *knows* is an introspective belief, and in this case *M* is being asked to introspect his perceptual belief/knowledge about the result of the experiment. This introspection is a belief about a belief.<sup>2</sup> For example, suppose *M* perceives a yellow daffodil in a green field. Then we would say that he has the occurrent perceptual belief that the flower is yellow. Such an occurrent belief would involve *M*’s visual cortex (Zeki 1993; Seymour et al. 2016). In this situation, upon being asked what color the flower is in front of him, *M* would presumably report, “The flower is yellow.” Now let’s ask *M* to report if he has a determinate result for his belief about the color of a flower in front of him. *M* might ask in this case, “Are you asking me what the color of the flower is?” To this we would answer, “No, we are asking you if you now have a determinate belief about the color of the flower in front of you. In particular, we are asking you to introspect your belief about the color of the flower in order to verify that you have a determinate belief about the flower’s color.” In this case, *M* will form an introspective belief—a belief about a belief—because he will need to check his current beliefs to verify that he has a belief about a yellow flower in front of him. This introspective belief is a belief that will have another belief as a content; and in particular, the content of this introspective belief will be *M*’s occurrent perceptual belief.

Note that the fine-grainedness of mental contents ensures that *M*’s belief, or introspective state, or experience that the flower is yellow will always be different from his belief, or introspective state, or experience that the flower is red (Tye 1995; Dretske 1995; Perry 1977; Putnam 1975; Frege 1892). And the same applies when *M* is observing the detector in the experiment: the fine-grainedness of mental contents ensures that *M*’s belief, or introspective state, or sensation that the electron is spin up will be different from his belief, or introspective state, or sensation that the electron is spin down.<sup>3</sup> This fine-grainedness is important: the intentional content of a belief about yellowness is different from the intentional content of a belief about redness.

<sup>2</sup> As knowledge states are typically taken to be some form of true belief, then *M*’s belief about this knowledge state is a belief about a belief; hence, an introspective belief. From here on for consistency, I will not refer to *M*’s perceptual belief of the experimental outcome as knowledge; rather, I will refer to any such occurrent perceptual state as a belief.

<sup>3</sup> Where the content in these cases includes, say, the position of a pointer on the measuring device; for example, pointing to one of the marks “+” or “-.”

And the intentional content of a belief that the dial points to “+” is different from the intentional content of a belief that the dial points to “-.” Such beliefs then will always differ. So, any sort of mental state with a content about spin up will differ from any mental state about spin down.<sup>4</sup> And, of course, it is the content of the state that allows us to call such states mental representations in the first place (Brentano 1874; Dretske 1988; Tye 1995). The content of a mental state also helps distinguish that type of state from other states, and in any causal theory of mental content, such states will be distinguished because of their differences in content, vehicle, and causal role (Skokowski 1999, 2018).

In order to have an introspective belief that can be used to report the occurrent belief being introspected, this introspective belief must be causally/physically connected with the occurrent perceptual belief through neural connections and neural firings. So, *M*’s introspective belief about a given perceptual belief will involve specific neural connections—as part of the introspective belief vehicle—to the specific portion of visual cortex that is involved with the exemplification of his perceptual belief, which itself involves a different set of neurons. In addition, when *M* has this introspective belief, then specific neural firings of the sort associated with the particular connections to this region of visual cortex will occur. That is, the action potentials that occur within the introspective belief in virtue of its connection with the color belief will be specific to that particular connection between the two sets of neurons. This introspective belief vehicle will therefore be made up of a specific set of neurons, which exemplify a certain pattern of neural firing (action potentials), when it occurs in *M*’s brain. Indeed, imaging studies appear to show that introspective beliefs occur in the pre-frontal cortex (Fleming et al., 2010).

And this analysis will apply for any of *M*’s mental states that are tasked with representing the content of his occurrent perceptual belief. Any such mental state will need to be causally connected to the occurrent perceptual belief, and the intentional content of the analyzing mental state will be fine-grained and hence *about* that occurrent perceptual belief and its content. That is, if *M* has a mental state *B* that is about his occurrent perceptual state and its content—where *B* is an introspective or some other self-analyzing state—then this mental state *B* will have a fine-grained content that is about

<sup>4</sup> Note that this difference in content holds whether the content is the position of a pointer towards either “+” or “-” or whether the content is actual color content “yellow” or “red.” The intentional content will be fine-grained in either case.

that occurrent perceptual belief and its content. Hence, any query to  $M$  about an occurrent perceptual belief state of his will depend for its answer on a mental state of his that represents that occurrent perceptual belief, and the intentional content of that representing mental state will have a fine-grainedness that tracks the fine-grainedness of the occurrent perceptual belief it represents.<sup>5</sup>

So, for example, the eigenstate corresponding to  $M$ 's introspection of his belief that the electron is  $x$ -spin up will contain his introspective state in the pre-frontal cortex as well as his perceptual belief in the visual cortex, and would therefore be written:

$$|\text{Introspect } x\text{-spin up}\rangle_{M-PF} |\uparrow\rangle_{M-VC} |\uparrow_x\rangle_S.$$

Or, more compactly,

$$|\text{Introspect}\uparrow\rangle_{M-PF} |\uparrow\rangle_{M-VC} |\uparrow_x\rangle_S,$$

where the subscripts " $M$ - $PF$ " stand for the introspective state in  $M$ 's pre-frontal cortex and the subscripts " $M$  - $VC$ " stand for the perceptual belief involving  $M$ 's visual cortex.

Let's apply this analysis to how the initial state of  $M$ 's brain, and the electron state he is observing, evolve from being "ready" to how things end up after the measurement. Before the measurement, we have:

$$|PF \text{ ready}\rangle_{M-PF} |VC \text{ ready}\rangle_{M-VC} (|\uparrow_x\rangle_S + |\downarrow_x\rangle_S),$$

where the first state is the "ready" state of the pre-frontal cortex, the second state is the "ready" state of the visual cortex, and the final state is the superposition of the  $x$ -spin directions up and down (that is, an electron initially in an eigenstate of  $z$ -spin expanded in the  $x$ -spin basis) for the electron that is about to be passed through the detector.

After the measurement, this state evolves into a superposition, call it  $|\psi\rangle$ , of the form:

$$\begin{aligned} |\psi\rangle = & 1/\sqrt{2}(|\text{Introspect}\uparrow\rangle_{M-PF} |\uparrow\rangle_{M-VC} |\uparrow_x\rangle_S \\ & + |\text{Introspect}\downarrow\rangle_{M-PF} |\downarrow\rangle_{M-VC} |\downarrow_x\rangle_S). \end{aligned}$$

<sup>5</sup> If the representing state did not have this fine-grainedness, then  $M$  would not be capable of answering queries about the content of the perceptual state in question.

Note that the first two eigenstates of both terms in the superposition represent different mental/belief states of  $M$ 's. As such, they will have different neural vehicles and contents from one another. In addition, were we to measure any of their neural properties (for example, which neurons, their action potentials, their connections, etc.), they would yield different eigenvalues. In the first term of the superposition, the first (introspective) state would give a measurement of a state of an introspection of a belief that the electron is spin up, and the second state would give a measurement of a perceptual belief that the electron is spin up. In the second term of the superposition, the first (introspective) state would give a measurement of a state of an introspection of a belief that the electron is spin down, and the second state would give a measurement of a perceptual belief that the electron is spin down. All these states differ in vehicle (the particular state in cortex) and content, which must be the case given the fine-grainedness of mental states.

Since each component of this superposition yields a different eigenvalue for its corresponding operator, then there will *not* be a common eigenvalue between the two components that would serve as an eigenvalue for an operator on the superposition taken as a whole. This is because of the fine-grainedness of mental states. Since each mental state in the superposition has its own vehicle and fine-grained intentional content, then each one will differ: each one will have a different eigenvalue from the other. This has to be the case for mental states, if they are to be individuated *qua* mental states. An introspective state about a belief that an electron is spin up is different from an introspective state about a belief that an electron is spin down. A perceptual belief that an electron is spin up is different from a perceptual belief that an electron is spin down. The vehicles differ: each state—introspective or perceptual—is composed of different neurons and action potentials, so the vehicles are exemplified as different physical states. And as Frege and others have taught us, the contents will also differ, as beliefs are known to be referentially opaque. For we know that even when  $m = n$ , a belief that  $m$  is  $P$  is not the same as a belief that  $n$  is  $Q$ , even when  $P$  and  $Q$  are co-extensional, and refer to the same things (Frege 1892; Putnam 1975; Dretske 1988; Tye 1995).

There are two important things to note here. First, the state of the system corresponding to  $M$ 's introspection of his belief that the electron is spin up will not be given by equation. Instead, the state of the system will contain  $M$ 's introspective state in the pre-frontal cortex as well as his perceptual belief involving the visual cortex, as spelled out above with state  $|\psi\rangle$ . Second, the

superpositional state  $|\psi\rangle$  as it stands does not yield the result that Barrett has claimed, which is that there is a measurable eigenvalue that is common to both components of the superposition. If there were such an eigenvalue, then there would be an observable property of  $M$  for the superposition of brain states he finds himself in according to the linear equations of motion.

Barrett's solution to this problem is to have  $M$  answer "Yes" to a question about his mental state. The question is for  $M$ , "Did you get some determinate result to your  $x$ -spin measurement, either  $x$ -spin up or  $x$ -spin down?" (Barrett 1999, p. 97). And so, " $M$  would report that he got a determinate  $x$ -spin result when he did not determinately get up and did not determinately get down" (p. 97).

Note that in order to answer this question,  $M$  will, in both components of the superposition, need to introspect his perceptual belief about the result of the experiment in order to evaluate the disjunction " $x$ -spin up or  $x$ -spin down." Since the perceptual beliefs in either component of the superposition are different, then the contents of these introspections in either component will also be different from each other. In the first component,  $M$ 's introspection of the perceptual belief that the  $x$ -spin is up would reveal the answer to the disjunction to be "up," whereas the second component,  $M$ 's introspection of the perceptual belief that the  $x$ -spin is down, would reveal the answer to the disjunction to be "down." Let's refer to the contents of these two introspective states as (UP  $\vee$  DOWN: UP) and (UP  $\vee$  DOWN: DOWN).  $M$  would need intentional contents with values like these in order to answer Barrett's question correctly on either component of the superposition. And let's say that, upon formulating these introspective beliefs,  $M$  can now answer the question as posed, "Did you get some determinate result to your  $x$ -spin measurement, either  $x$ -spin up or  $x$ -spin down?"

But note that there is a peculiarity about answering such a question. Answers to questions are formulated in a different part of the brain: Broca's area. And Broca's area is an area of the brain that is tasked with linguistic output—not with introspection. These linguistic outputs include unconscious grammatical processing and the signals required to form the mouth and tongue in a particular configuration, exhaling breath in a certain manner, opening and closing the nasal passages, and so forth.<sup>6</sup> Introspection

<sup>6</sup> Because introspection is conscious (Dretske 1995; Moore 1903), and linguistic processing is unconscious (Pinker 1994), the latter is not a candidate for introspective beliefs.

is tasked with producing beliefs about beliefs (Dretske 1995), whereas Broca's area is not: it is tasked with producing linguistic outputs (Pinker 1994, 1997).

The outputs of intentional mental states are not to be confused with mental states themselves. Consider the simple example of drinking a beer. I believe there's six-pack in the fridge and I desire a cold one. These mental states cause me to open the fridge door, grab a bottle, twist off the cap, and take a drink. The belief and desire are mental states with intentional contents, but the reaching, twisting, and drinking are outputs that are caused by these representational states. These outputs are not themselves representational states: that is, states with intentional contents from a function to represent properties in the environment, and executive capacities to cause action (Dretske 1988; Papineau 1987). They are instead *outputs* of representational states: causal consequences of mental states that *do* have the requisite executive capacities and intentional contents.

Barrett's question therefore is designed to detect a common measurable property in the  $M + S$  system by detecting a common output, and not by detecting a common property of the introspective representational eigenstates themselves. This is peculiar, because what is presumably at issue is the content of  $M$ 's own introspective and occurrent perceptual beliefs.

Note that Barrett is not asking  $M$  to introspect what spin result he perceived. The reason for this prescription is clear: The eigenvalue for asking  $M$ , "Do you introspect you are perceiving spin up?" will be different on both sides of the superposition, as it will if we decide instead to ask  $M$ , "Do you introspect you are perceiving spin down?," and so there will be no common eigenvalue for the superposition  $|\psi\rangle$  if either of these questions is asked. So a different question—a disjunctive one—must be asked. But that means that, rather than measuring  $M$ 's introspective states and their contents with an operator operating on them directly, we are instead being asked to measure a common output of those introspections. The focus has been shifted from introspection itself—and so a question about  $M$ 's beliefs—to a common *output* of introspective beliefs.

### 3. Linearity and Deception

The problem of shifting the focus in this way can be illustrated by an example. Note that if we can detect a common eigenvalue in the  $M$ -spin-detector superposition by means of a common output, rather than by the



representational state itself, then we should be able to detect a common eigenvalue in a *measuring device* by a common output as well. After all, what is important about an electron-spin measuring device is the representational state it ends up in: pointing to “+,” or pointing to “–,” for example. These pointer states are representational states with a content: they are *about* something, and crucially, they are *about* whether the measured electron is spin up or spin down. That’s what makes them representational in a way appropriate for measuring the *actual* spin of a particular electron.

Suppose, for example, that the needle on our electron-spin measuring device slides on the  $x$ -axis. It slides left for a spin-up electron and right for a spin-down electron. Now suppose we notice that whenever the measuring device registers the spin of an electron, there is a slight pressure wave of air in the  $y$ -axis normal to the  $x$ -axis. Due to turbulence effects, this wave is the same whether the electron is spin up or spin down. So we place a pressure detector on the  $y$ -axis that detects when such a pressure wave occurs.

Before we had this pressure detector attached, our superposition looked like:

$$|\Phi\rangle = 1/\sqrt{2}(|\uparrow\rangle_M |\uparrow_x\rangle_S + |\downarrow\rangle_M |\downarrow_x\rangle_S).$$

And after we attached the pressure detector, we have (where the subscript  $p$  refers to this detector):

$$|\Phi'\rangle = 1/\sqrt{2}(|\text{“y pressure”}\rangle_p |\uparrow\rangle_M |\uparrow_x\rangle_S + |\text{“y-pressure”}\rangle_p |\downarrow\rangle_M |\downarrow_x\rangle_S).$$

Call the magnitude of this pressure wave  $\lambda$ . Then we can see, by linearity, that a measurement of this pressure wave by an operator  $O$  is an observable property of the superpositional state as well as of each component of the superposition. So, by putting an electron in an eigenstate of  $z$ -spin through an  $x$ -spin detector, we can measure the observable property  $\lambda$ :

$$O|\Phi'\rangle = \lambda|\Phi'\rangle.$$

But notice that measuring a pressure wave like this—that is, an output of a representational/detecting state—does not mean that the detector is confused or deceived about what it’s detecting. It just means that the detector produces a measurable pressure wave regardless of whether it has collapsed to one component of the superposition or the other, or that, if Everett is correct, this observable will be measurable even in a superposition (by virtue

of linearity). A measurement of a  $y$ -pressure wave makes no claim about the final position of the pointer in the  $x$ -direction and hence the spin of the electron. This is explicitly spelled out in the state vector  $|\Phi'\rangle$ , where the two properties occupy different eigenstates, for example,  $|“y\text{-pressure}”>_p$  and  $|“x\text{-spin up}”>_m$ , that yield different physical properties for the device. Measurement of one property of the device does not mean deception about another property, as these are separate eigenstates with their own associated operators and eigenvalues. So the detector is not confused or deceived. There is just a common measurable output that occurs regardless of the QM interpretation.

And now we can say the same about the human observer  $M$ . When we consider the linguistic output from Broca’s area, then before the measurement, we have:

$$|B \text{ ready}\rangle_{M-B}|PF \text{ ready}\rangle_{M-PF}|VC \text{ ready}\rangle_{M-VC}(|\uparrow_x\rangle_S + |\downarrow_x\rangle_S),$$

where the first state is the “ready” state of  $M$ ’s Broca’s area, and the rest of the “ready” states are defined as before:  $M$ ’s pre-frontal cortex, his visual cortex, the “ready” state of the measuring device, and the final state is the superposition of up and down spin for the electron that is about to be passed through the detector. And again,  $M$ ’s eigenstates are designated by subscripts  $M-B$ ,  $M-PF$ , and  $M-VC$ .

After the measurement, this state evolves into a superposition of the form  $|\psi'\rangle$ , which is different from the superposition  $|\psi\rangle$  we considered earlier:

$$\begin{aligned} |\psi'\rangle = & 1/\sqrt{2}(|“Yes”\rangle_{M-B}|\text{Introspect}\uparrow\rangle_{M-PF}|\uparrow\rangle_{M-VC}|\uparrow_x\rangle_S \\ & + |“Yes”\rangle_{M-B}|\text{Introspect}\downarrow\rangle_{M-PF}|\downarrow\rangle_{M-VC}|\downarrow_x\rangle_S). \end{aligned}$$

Then we can see, by linearity, that a measurement of  $M$ ’s linguistic output will be an observable property of the superpositional state as well as of each component of the superposition. That is, when the superposition  $|\psi'\rangle$ , which includes  $M$ ’s states, is asked whether  $M$  has some definite belief in the way prescribed by Barrett, where this question is the operator  $O$ , then he will answer “Yes”:

$$O|\psi'\rangle = “Yes”|\psi'\rangle.$$

And this result, of course, is by virtue of this operator  $O$  operating on  $M$ ’s state  $|“Yes”>_{M-B}$ .

But this is like the detector example given immediately above. “Measuring” an answer like this—that is, an output of an introspective/representational state—does not mean that  $M$  is deceived about what he’s introspecting. It just means that  $M$  produces the answer “Yes” regardless of whether he has collapsed to one component of the superposition or the other, or that, if Everett is correct, this answer will be measurable even in a superposition (by virtue of linearity). A measurement of a spoken output “Yes” makes no claim about  $M$ ’s occurrent introspective state and his occurrent perceptual belief about the spin of the electron. This is explicitly spelled out in the state vector  $|\psi'\rangle$ , where the three properties are given by three different eigenstates  $|\text{“Yes”}\rangle_{M-B}$ ,  $|\text{Introspect } \uparrow\rangle_{M-PF}$ , and  $|\uparrow\rangle_{M-VC}$  of  $M$ ’s Measurement of one property of  $M$  does not mean deception about another property, as these are separate eigenstates with their own associated operators and eigenvalues. So when Barrett claims that  $M$  “would believe that he knows what the result is” (Barrett 1999, p. 98) (one of spin up or spin down) based on this spoken output, and that this belief is “false” (p. 98), we see that there is no basis for this claim, as  $M$  does not actually have the belief in question. That is, there is no single belief state that emerges from the superposition with the singular content being one of spin up or spin down. So  $M$  is not deceived. There is just a common linguistic output that occurs regardless of the QM interpretation.

In addition, it is important that deception is having a false belief: for  $X$  to be deceived about  $G$  is for  $X$  to believe  $G$  when  $G$  is not the case. We see that on either side of the superpositions involving  $M$ , there are only two beliefs: an introspective belief and a perceptual belief. As deception is a type of belief, these are the only candidates. However, none of these beliefs are instances of deception, because Barrett has stipulated beforehand that “ $M$  is a *good*  $x$ -spin observer in Everett’s sense (indeed, one might call him a perfect observer). . . .” (Barrett 1999, p. 96), so his beliefs within each component of the superposition—perceptual and introspective—about pointer position are accurate. This means any deception about the outcome must be at the level of speech output; that is, at the level of  $M$  saying, “Yes” about introspecting a definite belief about the spin of the electron. But speech output, as we have shown, is not an introspective state, and indeed, being an output rather than an intentional state, it is not a belief at all about the spin of the electron.

Barrett goes on to point out another peculiar aspect of the bare theory—that is, the nature of what he calls disjunctive experiences. In observing the experiment mentioned at the beginning of this paper, a proponent of

this theory “would not say that  $M$  would determinately believe that he had recorded  $x$ -spin up, nor would he say that he would believe that he had recorded  $x$ -spin down; rather, he would say that  $M$  would determinately believe that he had recorded  $x$ -spin up *or*  $x$ -spin down. One might call the experience leading to this disjunctive belief a disjunctive experience” (Barrett 1999, pp. 110, 111).

The problems with a claim like this from a bare theorist are twofold: first, in analyzing the superpositional states of the  $M + S$  system in any of the permutations we have considered, the belief states (either introspective or perceptual) are always on both sides of the superposition, with different vehicles, contents, and causal roles within their respective component of the superposition. That is, there is no single meta-belief formed with a disjunctive content like that proposed. The belief-chains, if you will, are formed on either side of the superposition, and are related to one another through their contents, neural connections, action potentials, etc. And neither of these chains can lead to such a disjunctive content, as each chain begins with a different spin for the particle, which is then perceived as that spin, then introspected as that spin, and so forth. The transparency of beliefs on either side of the superposition ensure that the contents of any introspective beliefs are of the singular contents within the respective component. Indeed, even the introspective “disjunctive” contents considered earlier contain components, “UP” on one side of the superposition and “DOWN” on the other side [(UP  $\vee$  DOWN: UP) and (UP  $\vee$  DOWN: DOWN)], that confirm which side of the superposition they reside in, so they are not pure disjunctions like the one given in the quote above. Which leads immediately into the second problem. As G. E. Moore painstakingly showed over a century ago, introspective states are *transparent*. That is, when one introspects another of one’s own mental states, one’s awareness is of the content of the mental state being introspected, not the mental state itself. Thus, any introspective states are drawn to the contents of the state being introspected: introspection therefore reveals no further phenomenal character or intentional content than the content of the sensation or belief being introspected (Moore 1903; Dretske 1995; Tye 1995). And the beliefs on both sides of the superposition contain contents specific to that side of the superposition (“+” or “-”), but never a content of a disjunctive belief with a content of the form “ $x$ -spin up *or*  $x$ -spin down.” And no mechanism has been provided by the bare theorist to show that such a belief has been formed. In order to do this, the bare theorist would need to explicate the vehicle, content, and causal role

(including its origin) for such a belief state. And that is a tall order that I don't believe the bare theorist has yet provided, because neither component of the superposition contains such a disjunctive content.

#### 4. Conclusion

We have seen that two claims of the bare theory about the observer *M* of a superposition fall short. First, when the observer *M* answers "Yes" to a question about his perceptual beliefs of the measurement, this does not imply that *M* would therefore falsely believe that he *knows* what the result of the measurement is. The reason, we have seen, is that *M* does not actually have the belief in question: there is no single belief state of *M*'s that emerges from the superposition with the singular content being one of spin up or spin down. Instead, as we have seen, there is a common linguistic output that occurs on both sides of the superposition, and such linguistic outputs of intentional states are not themselves intentional, and so are not themselves beliefs. We have also seen that *M* does not end up with a disjunctive belief about the experiment. The reason for this is that there is no mechanism for producing a meta-belief with such a content from the beliefs available to *M*. In both cases, these results were arrived at by carefully considering the vehicles, contents, and causal roles of *M*'s introspective and perceptual beliefs.

#### Acknowledgments

Thanks to John Perry, Reed Guy and Harvey Brown for very helpful discussions on the issues discussed here. All mistakes and misunderstandings are entirely my own.

#### References

- Barrett, J. (1999), *The Quantum Mechanics of Minds and Worlds*. Oxford: Oxford University Press.
- Brentano, F. (1874), *Psychologie vom Empirischen Standpunkt*. Leipzig.
- Dretske, F. (1988), *Explaining Behavior*. Cambridge, MA: MIT Press.
- Dretske, F. (1995), *Naturalizing the Mind*. Cambridge, MA: MIT Press.
- Fleming, S., et al. (2010), "Relating introspective accuracy to individual differences in brain structure." *Science*, 329(5998): 1541–1543.

- Frege, G. (1892), "On sense and reference" [Über Sinn und Bedeutung]. *Zeitschrift für Philosophie und philosophische Kritik*, 100: 25–50.
- Lee, T. S., et al. (1998), "The role of the primary visual cortex in higher level vision." *Vision Research*, 38: 2429–2454.
- Moore, G. E. (1903), "The refutation of idealism." *Mind*, 12: 433–453.
- Papineau, D. (1987), *Reality and Representation*. Oxford: Blackwell.
- Perry, J. (1977), "Frege on demonstratives." *Philosophical Review*, 86: 474–497.
- Pinker, S. (1994), *The Language Instinct*. New York: HarperCollins.
- Pinker, S. (1997), *How the Mind Works*. New York: W. W. Norton.
- Putnam, H. (1975), "The meaning of 'meaning.'" *Minnesota Studies in the Philosophy of Science*, 7: 131–193.
- Seymour, K. J., et al. (2016), "The representation of color across the human visual cortex: Distinguishing chromatic signals contributing to object form versus surface color." *Cerebral Cortex*, 26: 1997–2005.
- Skokowski, P. (1999), "Information, belief, and causal role." In Moss et al. (ed.), *Logic, Language and Computation*. Stanford, CA: CSLI Press.
- Skokowski, P. (2018), "Temperature, color and the brain: An externalist response to the knowledge argument." *Review of Philosophy and Psychology*, 9(2): 287–299.
- Tye, M. (1995), *Ten Problems of Consciousness*. Cambridge, MA: MIT Press.
- Zeki, S. (1993), *A Vision of the Brain*. Oxford: Blackwell.