

A New Theory of Free Will

Marcus Arvan
University of Tampa
marvan@ut.edu

Abstract: This paper shows that the conjunction of several live philosophical and scientific hypotheses – including the holographic principle and multiverse theory in quantum physics, and eternalism and mind-body dualism in philosophy – jointly imply an audacious new theory of free will. This new theory, "Libertarian Compatibilism", holds that the physical world is an eternally existing array of two-dimensional information – a vast number of possible pasts, presents, and futures – and the mind a nonphysical entity or set of properties that "read" that physical information off to subjective conscious awareness (in much the same way that a song written on an ordinary compact-disc is only *played* when read by an outside medium, i.e. a CD-player). According to this theory, every possible physical "timeline" in the multiverse may be *fully physically deterministic or physically-causally closed* but each person's consciousness still entirely free to choose, *ex nihilo*, outside of the physical order, which physically-closed timeline is experienced by conscious observers.

Although Libertarian Compatibilism is admittedly fantastic, I show that it not only follows from several live scientific and philosophical hypotheses, I also show that it (A) is a far more explanatorily powerful model of quantum mechanics than more traditional interpretations (e.g. the Copenhagen, Everett, and Bohmian interpretations), (B) makes determinate, testable empirical predictions in quantum theory, and finally, (C) predicts and explains the very existence of a number of philosophical debates and positions in the philosophy of mind, time, personal identity, and free will. First, I show that whereas traditional interpretations of quantum mechanics are all philosophically problematic and roughly as ontologically "extravagant" as Libertarian Compatibilism – in that they all posit "unseen" processes – Libertarian Compatibilism is nearly identical in structure to the only working simulation that human beings have ever constructed capable of reproducing (and so explaining) *every general feature of quantum mechanics we perceive*: namely, massive-multiplayer-online-roleplaying videogames (or MMORPGs). Although I am not the first to suggest that our world is akin to a computer simulation, I show that existing MMORPGs (online simulations we have already created) actually reproduce every general feature of quantum mechanics within their simulated-world reference-frames. Second, I show that existing MMORPGs also replicate (and so explain) many philosophical problems we face in the philosophy of mind, time, personal identity, and free will – all while conforming to the Libertarian Compatibilist model of reality.

I conclude, as such, that as fantastic and metaphysically extravagant as Libertarian Compatibilism may initially seem, it may well be true. It explains a number of features of our reality that no other physical or metaphysical theory does.

A New Theory of Free Will

5.641 The philosophical self is not the human being, not the human body, or the human soul, with which psychology deals, but rather the metaphysical subject, the limit of the world – not a part of it.

– Wittgenstein, *Tractatus Logico-Philosophicus*

This paper shows that several serious philosophical and scientific hypotheses – hypotheses that have received, and continue to receive, serious philosophical and scientific defense, and which are serious epistemic possibilities, given our present evidence – jointly imply an audacious new theory of free will. The hypotheses to be discussed are:

Eternalism: the hypothesis that past, present, and future objects and properties all exist “timelessly.”¹

The Multiverse Hypothesis: the hypothesis that the observable universe is merely a small part of a multiverse which contains a vast, possibly infinite array of alternative pasts, presents, and futures.²

I am indebted to an anonymous reviewer, whose encouraging and challenging comments helped to inspire this paper to become what it is. I would also like to thank another anonymous reviewer and Douglas Lackey at *The Philosophical Forum*.

¹ Philosophical proponents of eternalism include Adolf Grünbaum “Relativity and the Atomicity of Becoming,” *Review of Metaphysics* (1950-51): 143-186; J.J.C. Smart, “Spatialising Time,” *Mind* 64 (1955): 239-241; W.V.O. Quine, *Word and Object* (MIT Press, 1960); Paul Horwich, *Asymmetries in Time* (MIT Press, 1987); Ted Sider, “Presentism and Ontological Commitment,” *Journal of Philosophy* 96, (1999): 325-347; and Ted Sider, *Four-Dimensionalism: An Ontology of Persistence and Time* (Oxford University Press, 2001). Many physicists also defend eternalism. See e.g., Stephen Hawking and Leonard Mlodinow, *The Grand Design* (Bantam Press, 2010); Jarek Duda, “Four-dimensional understanding of quantum mechanics.” *arXiv:0910.2724* (2009); Petkov Vesselin, “Is There an Alternative to the Block Universe View?,” *PhilSci Archive*, (2005), accessed on 10/18/2012 at philsci-archive.pitt.edu/2408/ -- and many others. Also see Steven F. Savitt, “Presentism and Eternalism in Perspective”, (unpublished manuscript), accessed on 10/18/2012 at philsci-archive.pitt.edu/1788/, who defends the intriguing view that presentism and eternalism are not contradictory but rather complementary elements of a naturalistic philosophy of time. Finally, see Hrvoje Nikolic, “Block time: why many physicists still don’t accept it?”, (unpublished manuscript), accessed on 10/18/2012 at fqxi.org/data/essay-contest-files/Nikolic_FQXi_time.pdf, who defends a similar dualistic theory of time according to which eternalism is true of “physical time” but presentism is true of “conscious time.” The present paper lends added support to these dualist theories of time.

The Holographic Principle: the hypothesis that in order to unify quantum mechanics and general relativity, the universe must be understood as comprised by information “written” on the cosmological horizon.³

Mind-body Dualism: the hypothesis that the mind, or at least certain properties of it, are non-physical.⁴

Subjectivity About the Flow of Time: the hypothesis that the passing of time is not in the objective physical world but rather within us (i.e. within consciousness).⁵

² The many-worlds interpretation of quantum mechanics, first defended by Everett, Hugh, "Relative state' formulation of quantum mechanics". *Reviews of Modern Physics* 29, (1957): 454–462, and presently one of several mainstream interpretations of quantum mechanics, posits a concrete universe for every possible state of the quantum wave-function.

³ The Holographic Principle, first proposed by Gerard 't Hooft, "Dimensional Reduction in Quantum Gravity", *arXiv:gr-qc/9310026*, (1993), has since been given a precise string theory formulation by Leonard Susskind, "The World as a Hologram", *Journal of Mathematical Physics*, 36 (1995): 6377–6396. Also see Raphael Bousso, "The holographic principle", *Reviews of Modern Physics* 74, (2002): 825–874. The Holographic Principle's only unique empirical prediction to date – the prediction that the holographic principle may imply quantum fluctuations in spatial prediction leading to observable background noise in gravitational wave detectors (see Craig J. Hogan, "Measurement of quantum fluctuations in geometry". *Physical Review D* 77 (10), (2008): 104031) – may have already been verified. See Marcus Chown, "Our world may be a giant hologram", *New Scientist* (2009): 2691, for a summary of the experimental results so far.

⁴ Dualism has had many notable contemporary defenders. See e.g., David Chalmers *The Conscious Mind* (New York: Oxford University Press, 1996); John Foster, *The Immaterial Self*. (London: Routledge, 1991); Brie Gertler, "A Defense of the Knowledge Argument." *Philosophical Studies* 93 (3) (1999):317-336; Brie Gertler, "Consciousness and Qualia Cannot Be Reduced." In Robert J. Stainton (ed.), *Contemporary Debates in Cognitive Science (Contemporary Debates in Philosophy)*, (Blackwell, 2006); Frank Jackson, "What Mary didn't Know", *Journal of Philosophy*, 83, 5 (1986): 291-295; Saul Kripke, *Naming and Necessity* (Oxford: Blackwell, 1980); Thomas Nagel, "What Is it Like to Be a Bat?", *Philosophical Review*, Vol. 83, No. 4, (1974): 435-50; Bertrand Russell, *The Analysis of Mind* (London: G. Allen & Unwin; New York, Macmillan, 1921); Gregg Rosenberg, *A Place for Consciousness: Probing the Deep Structure of the Natural World* (Oxford University Press, 2004); and Marcus Arvan, "A Simple Proof of Mind-Body Dualism", (unpublished manuscript), accessed on 10/18/2012 at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1636024. Also see William G. Lycan, "Giving Dualism Its Due." *Australasian Journal of Philosophy*, Vol. 87, No. 4 (2009): 551-563, who is a physicalist but has recently argued that, "no convincing case has been made against dualism, and...standard objections to it can be credibly answered." (551)

⁵ This is perhaps the most unpopular hypothesis that my argument relies upon. However, it does have serious contemporary proponents (e.g. Huw Price, "The Flow of Time", in Craig Callender (ed.), *The Oxford Handbook of Time* (Oxford: Oxford University Press, 2011), as well as famous historical proponents. Kant is the most notable proponent of the subjectivity of time (see Immanuel Kant, *Critique of Pure Reason*, [1781] translated/reprinted in Guyer, P., and Wood, A., eds., (1998), *Critique of Pure Reason* (Cambridge: Cambridge University Press), though others have defended the view too, including Henri Bergson, *Time and Free Will: An Essay on the Immediate Data of Consciousness*, [1889]

The Further Fact Theory of Personal Identity: the hypothesis that personal identity is a brute, simple fact that cannot be reduced to any sort of physical or psychological relation.⁶

Single Commonly-Experienced (or“Actualized”) Timeline: the hypothesis that only one physical universe – our Universe – is experienced by conscious observers.⁷

tr., F.L. Pogson (Montana: Kessinger Publishing Company, 1910); Georg Wilhelm Hegel, *Phenomenology of Spirit*, [1807] trans. A. V. Miller (Oxford: Oxford University Press, 1977); and Nikolic (unpublished). Finally, I give a new argument for subjectivity about time here, in §1.5. My argument is that if existing arguments for mind-body dualism succeed, those arguments also establish dualism regarding time. For, I argue, the flow of time and the “now-ness” of the present moment are no more explainable in physical terms than any other “qualia” we experience. Thus, if qualia are non-physical (and there are strong though debatable arguments that they are), then, since the passage of time and “now-ness” are *themselves* qualia, a mind-body dualist should accept (a) eternalism for the physical universe, but (b) a presentist view of time for the *subjective, conscious flow* of time. Finally, as we will see, this position has the added upshot of synthesizing three independently plausible but otherwise mutually inconsistent theories of time: presentism, eternalism, and the “moving spotlight” view.

⁶ The further-fact theory is not very popular today. However, it has had many contemporary defenders – see e.g. Roderick Chisholm, “The Loose and Popular and the Strict and Philosophical Senses of Identity,” in *Perception and Personal Identity*, ed. Norman Care and Robert Grimm, 82-106 (Cleveland: Press of Case Western University, 1969); Geoffrey Madell, *The Identity of the Self* (Edinburgh: Edinburgh University Press, 1981); Richard Swinburne, “Personal Identity: The Dualist Theory”, in *Personal Identity*, eds. Richard Swinburne and Sidney Shoemaker, 1-66 (Oxford: Blackwell, 1984); E.J. Lowe, *Subjects of Experience* (Cambridge: Cambridge University Press, 1996); Trenton Merricks, “There Are No Criteria for Identity Over Time”, *Noûs*, 32 (1998): 106-24; and most recently Dilip Ninan, “Persistence and the First-Person Perspective”, *Philosophical Review*, 188, 4 (2009): 425-464. It has also had numerous proponents throughout the history of philosophy, including Joseph Butler, *Of Personal Identity. First appendix to The Analogy of Religion* (1736), Reprinted in ed., John Perry, *Personal Identity* (Berkeley: University of California Press, 1975), and Thomas Reid, *Essays on the Intellectual Powers of Man*, (1785), Reprinted in Derek Brookes (ed.), (University Park: Pennsylvania State University Press, 2002). I would also suggest adding Kant and Wittgenstein (1922) to the list, given their views that we are in some way transcendental subjects of experience (see Kant [1781]; Immanuel Kant, *Critique of Practical Reason*, [1788], translated/reprinted in Gregor, M., (ed.) (Cambridge University Press, 1996); Immanuel Kant, *Groundwork of the Metaphysics of Morals*, [1785], translated/reprinted in Gregor, M., (ed.), 1996, *Practical Philosophy* (Cambridge: Cambridge University Press): section III; and Ludwig Wittgenstein, *Tractatus Logico-Philosophicus*, translated by C. K. Ogden (Routledge&Kegan Paul, 1922). Finally, as we will see in §1.6, the most common argument against the further-fact theory – the argument that it is incompatible with physicalism (see Ninan 2009) – is an argument that any dualist (of which I am one) should reject.

⁷ As I explain in §1.7., this hypothesis is the default position in philosophy. Insofar as this is the only universe that any of us have (to our knowledge) ever experienced, Occam’s Razor (i.e. ontological parsimony) requires us to treat this hypothesis as a starting assumption – one that we should reject if, and only if, doing so is theoretically necessary to explain reality.

§1 of this paper briefly summarizes each hypothesis, as well as the evidence in favor of each. §2 then shows that the conjunction of these hypotheses implies the following argument for “Libertarian Compatibilism”:

The Argument for Libertarian Compatibilism

- A. There is presently some positive evidence that the physical universe is a timelessly existing array of information comprising a vast, possibly infinite variety of branching “time-lines” (in much the same way that different songs written on the surface of an ordinary compact-disc are just a variety of different ordered series’ of digital information encoded upon the disc).
- B. There is presently some positive evidence that each person’s consciousness is a non-physical entity or set of non-physical properties that:
 - i. “Flow(s) over” the eternally existing world of physical information, “reading” that information off to each person’s consciousness (in much the same way that an ordinary compact-disc player’s laser is an entity outside of the disc that flows over the surface of the disc, playing the information to observers),
 - ii. Possesses the genuine capacity to choose, *ex nihilo*, outside of the physical order, which physical “timeline” they experience in subjective awareness, such that,
 - iii. The joint choices of all conscious observers “collapse” all possible “timelines” at every instant to a single, “actualized” reality that all conscious observers experience in tandem. (See Figure 1 for illustration).

- C. If (A) and (B) are true, then there is presently some positive evidence in favor of Libertarian Compatibilism: the theory that genuine libertarian free will (the ability of conscious observers to choose their actions *ex nihilo*, outside of the physical-causal order) *coexists* with a completely deterministic or causally-closed physical order. Free will consists of each person's non-physical consciousness:
- i. "Flowing over" the eternally existing physical world (much as a laser on a compact disc reads physical information encoded on the disc),
 - ii. Choosing, *ex nihilo*, which particular "path" of possible futures one experiences in a purely subjective experience of time's passage, such that,
 - iii. The joint choices of all conscious observers "collapse" possible paths through the multiverse at every given instant to a single, intersubjectively experienced (or "actualized") reality.
- D. Thus (from premises A-C), there is presently some positive evidence in favor of Libertarian Compatibilism: the theory that free will consists of each person's non-physical consciousness "flowing over" the eternally existing physical world, choosing, *ex nihilo*, which particular "path" of possible futures one experiences in a purely subjective experience of time's passage, such that the joint choices of all conscious observers "collapse" possible paths through the multiverse at every given instant to a single, intersubjectively experienced (or "actualized") reality.

Figure 1.
Libertarian Compatibilism
(Modeled for Two Conscious Individuals, at Times t and $t+1$)

Grid: array of physical “timelines” through the eternally-existing multiverse

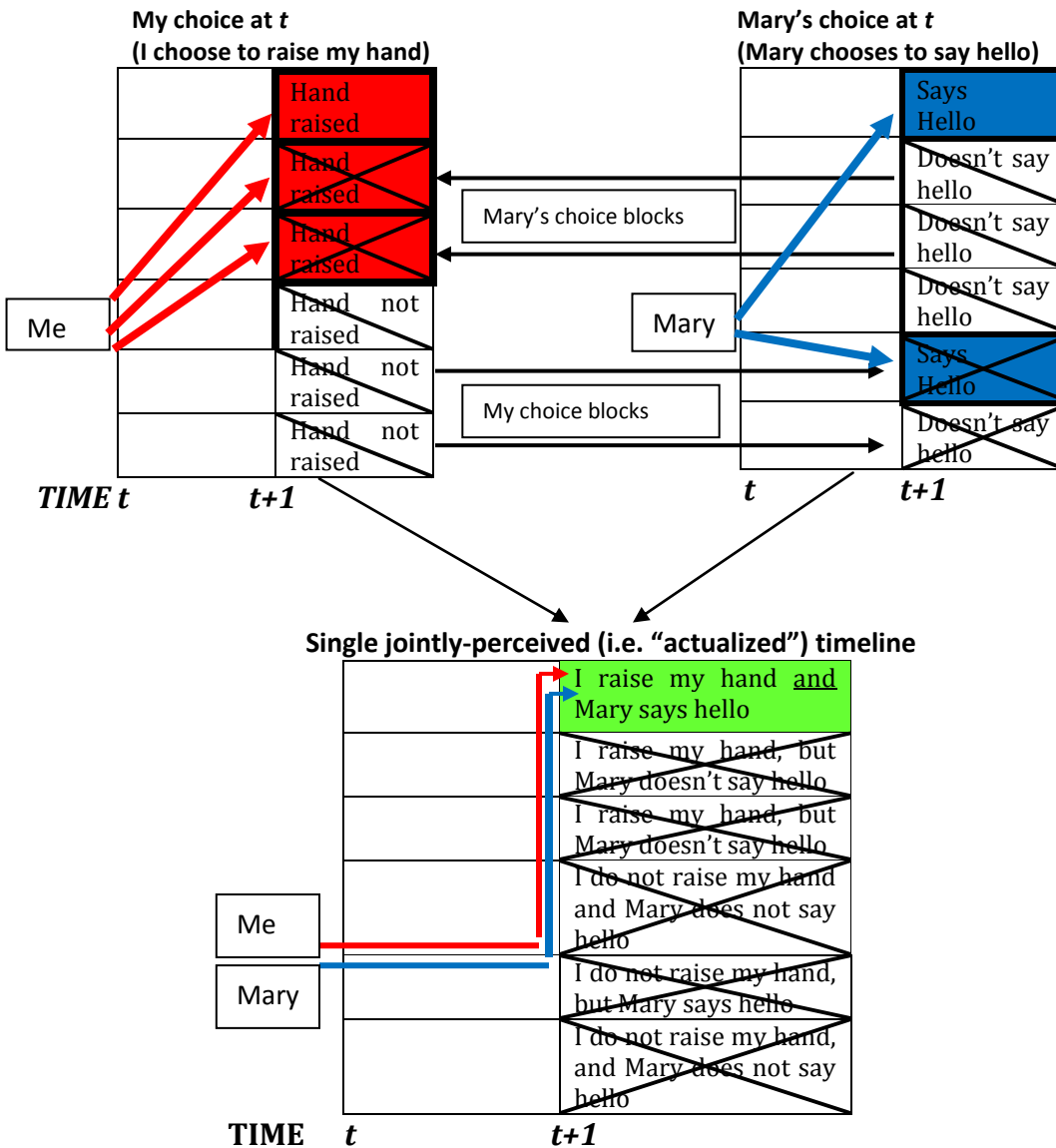
Red arrow: My consciousness

Blue arrow: Mary’s consciousness

Red boxes: multiverse “timelines” in which my conscious choice at t is realized at $t+1$ (I raise my hand)

Blue boxes: multiverse “paths” in which Mary’s choice at t is realized at $t+1$ (she says hello)

Green box: “quantum collapse” to single, commonly experienced event



Forthcoming in *The Philosophical Forum*

This is a fantastic theory, to be sure. However, it is not mere fancy. §3 of this paper shows that existing interpretations of quantum mechanics – e.g. the Copenhagen Interpretation, Everett interpretation, and Bohmian Mechanics – are all roughly as “ontologically extravagant” as Libertarian Compatibilism, insofar as every such interpretation posits the existence of objects and properties that are unobservable in principle. §4 then shows that although Libertarian Compatibilism posits unobservable entities, it does not invoke these entities *unnecessarily* (i.e. it does not violate Occam’s Razor). It is shown that the entities and properties Libertarian Compatibilism invokes are actually *necessary* to explain many of the “appearances” of our world: everything from quantum-mechanics to philosophical problems of mind, time, personal identity, and free will. Indeed, §4 shows that Libertarian Compatibilism in fact respects Occam’s Razor *better* than rival interpretations of quantum mechanics because it corresponds to the *only working simulation* that human beings have ever constructed capable of replicating every general feature of quantum mechanics, as well as a vast array of philosophical problems in the philosophy of mind, time, personal identity, and free will – namely, massive-multiplayer-online-roleplaying videogames (or MMORPGs). Although I am not the first to suggest that our world is akin to a computer simulation, I show that *already-existing* MMORPGs – videogames such as “The Sims” and “Halo” – replicate, *within their simulated world reference-frames*, everything from (a) quantum indeterminacy, to (b) problems of quantum measurement, to (c) the appearance of physical-causal closure, to (d) the appearance of an “explanatory gap” between the physical world and the subjective perspectives of observers, to (e) the appearance of the passage of time (as well as time’s apparent “unreality”), to (f) the thought that

personal identity is something more than physical or psychology continuity – and finally, (g) the problem of free will. Finally, I use MMORPGs to illustrate that unlike many existing interpretations of quantum mechanics, which appear to make no testable empirical predictions about the world, Libertarian Compatibilism makes several *determinate* empirical predictions, and is thus amenable to verification and falsification.

Because Libertarian Compatibilism both (I) follows from several live scientific and philosophical hypotheses, and – by analogy to MMORPGs – (II) explains far more about our reality (i.e. about quantum mechanics, and philosophy) than any existing physical or philosophical theory, I submit that the theory is not as far-fetched as it may initially seem, and indeed, may well be true. Libertarian Compatibilism not only explains the very nature of the quantum-physical world; it also explains the very existence of an entire swath of philosophical problems in the philosophy of mind, the philosophy of time, and personal identity. Finally, it is the only theory of free will to date capable of accomplishing a task which has traditionally seemed impossible: reconciling the apparent causal closure of the physical world – the appearance of every physical event as having a full physical explanation – with genuine libertarian free will, i.e. the capacity to make free choices in *no* way enacted by physical laws or processes.

§1. Evidence for the Six Hypotheses

One or more of the hypotheses I will now discuss may well turn out to be false. At present, however, there is evidence favoring each of them. Let me briefly explain each hypothesis, as well as the evidence in favor of each.

§1.1. Evidence for Eternalism

Eternalism is the hypothesis that, “objects from both the past and the future exist just as much as present objects.”⁸ There are, broadly speaking, two types of arguments in favor of eternalism. The first type of argument is *a priori*. The argument is roughly as follows:

1. There are truths about the past, present, and future.
2. There are truths about the past, present, and future if and only if objects and properties in the past, present, and future all exist eternally (or “timelessly”).
3. Thus, objects and properties – past, present, and future – exist eternally.⁹

Here is a simple illustration. Dinosaurs once existed. That proposition is true. But in order for that proposition to be true, there must *be* dinosaurs that once existed. But that, in turn, seems tantamount to saying that there *is* a previous time at which dinosaurs exist. So, it seems, past objects and properties exist. *Mutatis mutandis* for the future. Thus, past, present, and future objects and properties all exist “timelessly.”

Now, there are many ways to challenge this type of argument. One possibility is to deny that propositions about the future are literally true.¹⁰ Another popular option is to attempt to defend a *tensed semantics* that can account for the truth-value of past- and future-directed propositions without positing existent past or future objects (this view, if it could be sustained, would obviously be favored by Occam’s Razor, or ontological

⁸ Ned Markosian, “Eternalism”, *Stanford Encyclopedia of Philosophy*, (2007), accessed on October 18, 2012 at <http://plato.stanford.edu/entries/time/>: §6.

⁹ See e.g. Sider (1999, 2001). Cf. John Ellis McTaggart, “The Unreality of Time”, *Mind: A Quarterly Review of Psychology and Philosophy*, 17 (1908): 456-473. Markosian (2008) gives a similar argument for Fatalism, the view that all future events are unavoidable.

¹⁰ Markosian (2008): §1. Also see Aristotle’s problem of future contingents in *On Interpretation*: chapter 9.

simplicity).¹¹ It is not our task to evaluate these or other objections here (though eternalists tend to regard both objections as incoherent¹²). It suffices for our purposes that defenders of eternalism have rigorously replied to all of these objections, and that eternalism is a serious epistemic possibility defended today by many metaphysicians.

The second type argument for eternalism is empirical in nature. For example, many have argued that Einstein's theory of special relativity entails eternalism.¹³ Because special relativity denies the reality of objective simultaneity – all “simultaneous” events are only simultaneous-relative-to-some-frame-of-reference – special relativity entails that there can be no physical basis for picking out a unique set of events happening simultaneously in “the present.” Events that appear to be in the “past” from one frame of reference may appear to be in the present from another frame of reference, and in the “future” from still another frame of reference. Finally, insofar as special relativity arguably entails that all reference-frames are equally “valid” – it is possible, at least in principle, observe the world from *any*¹⁴ frame-of-reference in space-

¹¹ See Markosian (2008): §1. Also see Sider (1999, 2001).

¹² See e.g., McTaggart (1908), Sider (1999, 2009).

¹³ See Craig Callender, “Shedding Light on Time,” *Philosophy of Science* (Proceedings), LXVII (2000): S587-S599; William Godfrey-Smith, “Special Relativity and the Present,” *Philosophical Studies*, XXXVI (1979): 233-244; Nicholas Maxwell, “Are Probabilism and Special Relativity Incompatible?,” *Philosophy of Science* 52 (1) (1985): 23-43; C.W. Rietdijk, “A Rigorous Proof of Determinism Derived from the Special Theory of Relativity,” *Philosophy of Science*, XXXIII (1966): 341-344; C.W. Rietdijk, “Special Relativity and Determinism,” *Philosophy of Science*, XLIII (1976): 598-609; Steven Savitt, “The Replacement of Time,” *Australasian Journal of Philosophy*, LXXII (1994): 463-474; Steven Savitt, “There's No Time Like the Present (in Minkowski Spacetime),” *Philosophy of Science* (Proceedings), LXVII (2000): S563-S574; Simon Saunders, “How Relativity Contradicts Presentism”, in C. Callender, ed., *Time, Reality, and Experience* (Cambridge University Press): 277-92; Lawrence Sklar, *Space, Time and Spacetime* (Berkeley, CA: University of California Press, 1974); Lawrence Sklar, “Time, Reality, and Relativity,” in Richard Healey, ed., *Reduction, Time, and Reality* (New York: Cambridge University Press, 1981): 129-142; Howard Stein, Howard, “On Einstein-Minkowski Space-Time,” *The Journal of Philosophy*, LXV (1968): 5-23; Howard Stein, “A Note on Time and Relativity Theory,” *The Journal of Philosophy*, LXVII (1970): 289-294; and Howard Stein, “On Relativity Theory and Openness of the Future,” *Philosophy of Science*, LVIII (1991): 147-167.

¹⁴ Although Einstein asserted that the speed of light is in fact the Universe's “speed limit”, Einstein's theories of relativity strictly entail that *if* faster-than-light (FTL) travel is possible, then time-travel

time – the past, present, and future in some sense must exist (since there is always a reference-frame from which each can be viewed as “present”).

As with the above *a priori* arguments for eternalism, there is a great deal of debate about these empirical arguments. Some deny that special relativity entails eternalism.¹⁵ However, again, it is not our task here to enter these debates. It suffices for our purposes that these are all active debates, and thus, that there is presently some serious philosophical and scientific evidence in favor of eternalism.

§1.2. Evidence for the Multiverse Hypothesis

The multiverse hypothesis – that the observable universe is merely a small part of a much larger “multiverse” containing an infinite array of pasts, presents, and futures – is very popular today in quantum physics and cosmology. It has been defended by many noted physicists, including Hugh Everett, Max Tegmark, Michio Kaku, and others, both as (a) an implication of observations of the universe’s expansion (the so-called eternal inflation hypothesis¹⁶), as well as (b) an implication of quantum-mechanics (viz. the Everett/many-worlds interpretation).¹⁷ Thus, though it may eventually be discovered

into the distant past – i.e. travel to space-time reference frames in the past *relative to our reference frame* – is possible. Since it is presently an open empirical question whether anything can travel faster than light, it is unclear whether it is in fact *physically* possible to observe the world from every possible physical frame-of-reference. Some frames-of-references (those prior to our present “light-cone”) may be physically inaccessible. Still, the point is that *even if such reference-frames are not physically (or nomologically) possible to adopt*, they are strictly *metaphysically possible* to adopt – and that is all that is needed for the sort of argument for eternalism in this section to go through.

¹⁵ See e.g. Thomas Crisp, “Presentism, Eternalism, and Relativity in Physics”, in *Einstein, Relativity, and Absolute Simultaneity*, eds. William Lane Craig and Quentin Smith (Routledge 2007); and Mark Hinchliff “A Defense of Presentism in a Relativistic Setting,” *Philosophy of Science* (Proceedings) LXVII (2000): S575-S586.

¹⁶ See Alan Guth, “Eternal Inflation and Its Implications”, *Journal of Physics A: Mathematical and Theoretical* (2007), and Andrei Linde, “Eternally Existing Self-Reproducing Chaotic Inflationary Universe”, *Physics Letters B*, 175 (4) (1986): 395-400.

¹⁷ See e.g. Everett (1957), Max Tegmark, “The Mathematical Universe”, *Foundations of Physics*, 38 (2008): 101-150; and Michio Kaku, *Parallel Worlds: A Journey Through Creation, Higher Dimensions, and the Future of the Cosmos* (New York: Doubleday, 2004).

to be false, it seems safe to consider the Multiverse Hypothesis a serious epistemic possibility at present: a live hypothesis with some real evidence in its favor.

§1.3. Evidence for the Holographic Principle

Reconciling quantum mechanics with Einstein's theory of general relativity is one of the biggest problems today in theoretical physics. Quantum mechanics describes the world of the very small (the "micro world") very well, and seems to describe everything except for gravity, which has only been observed to emerge at larger, "macro" levels. Gravity and its effects at these macro levels are, in contrast, described by Einstein's theory of general relativity. Yet, quantum mechanics and Einstein's theory appear to be inconsistent. General relativity is a *classical* mechanical theory: that is, it models real-world objects as comprised by *point-particles with determinate properties* (i.e. determinate positions, velocities, etc). Quantum-mechanics, on the other hand, is a *non-classical* mechanical theory: it models real world objects as comprised by particles represented by the quantum wave-function, which ascribes only *probabilistic* properties to fundamental particles. Unless quantum mechanics and the theories of relativity somehow govern different realms altogether (i.e. macro- and micro-physical reality are governed by entirely distinct sets of laws), or there is a more powerful physical theory to (somehow) unify the two theories (i.e. a "grand unified theory"), at least one, if not both, of the two theories must be false.

Interestingly, a number of theoretical physicists have recently argued that quantum mechanics *can* make sense of gravity – but if, and only if, the universe is understood as fundamentally comprised by information "written" on a surface (the

cosmological horizon).¹⁸ According to this new theory – the Holographic Principle – gravity and time are both, strictly speaking, *holograms* that emerge from information written on the Universe’s cosmological horizon. Now, of course, this argument for the Holographic Principle – that its ability to unify general relativity and quantum mechanics is positive evidence for it – assumes that the two theories ought to be unified. But one can surely question this assumption. *Must* there be a unified physical theory? Why not wait for more data before working with empirically unverified principles or theories? These are fair questions. However, they should not prevent us from moving forward with our present project for two reasons. First, setting aside widely-discussed Humean worries about induction in general (which we cannot begin to debate here), we have a great deal of general inductive evidence in support of theoretical unification in the physical sciences. Scientific progress in the physical sciences has largely been a slow (and not always linear) march toward greater unification (e.g. we now understand life in terms of DNA and other biological processes, DNA in terms of molecules, molecules in terms of atoms, etc.). Accordingly, if the Holographic Principle is *a* way (the only way?) to unify general relativity and quantum mechanics, then we have significant inductive, empirical evidence – again, the incremental success of greater unification in physical sciences as a whole – in its favor. Second, even though the Holographic Principle itself is presently unverified by experimental evidence (though its first prediction may have already have been confirmed)¹⁹, this should not prevent us from investigating its philosophical and physical implications. It is, after all, common practice in philosophy and science to

¹⁸ Again, see ‘tHooft (1993), Susskind (1995), and Bousso (2002).

¹⁹ Again, see Hogan’s (2008) prediction and Chown (2009) for a summary of evidence so far compiled.

explore the implications of as-yet unverified models (good examples include supersymmetry in quantum physics, theories of causation in metaphysics, etc.).

In short, although the Holographic Principle is presently unverified, there are plenty of reasons to consider it a serious epistemic hypothesis, and to investigate its physical and philosophical implications.

§1.4. Evidence for Mind-Body Dualism

The mind-body debate has raged for millennia. Many philosophers are physicalists.²⁰ Others are anti-physicalists.²¹ Others still are unconvinced either way.²² Accordingly, the true nature of the mind clearly remains an open epistemic question. In particular, there are serious arguments for mind-body dualism: the idea that phenomenal consciousness is not identical to, and cannot be reduced (without remainder) to, physical objects, properties, or functional states. Frank Jackson's "knowledge argument", for example, is still taken very seriously by both proponents and opponents of dualism.²³ David Chalmers' "conceivability" argument (otherwise known as his "zombie" argument) is also quite influential.²⁴ Finally, in my opinion, the best argument for mind-body dualism is the one given by Russell, Rosenberg, and a few others.²⁵ That argument holds that science describes a world of relational properties – that is, how

²⁰ See e.g. D.M. Armstrong, *A Materialist Theory of the Mind* (London: Routledge and Kegan Paul, 1968); D.M. Armstrong, "The Nature of Mind." In *The Nature of Mind and Other Essays* (University of Queensland Press, 1980); Patricia Churchland, *Neurophilosophy: Toward a Unified Science of the Mind-Brain* (MIT Press, 1986); Daniel C. Dennett, *Consciousness Explained* (Back Bay Books, 1992); David K. Lewis "An Argument for the Identity Theory." *Journal of Philosophy* 63 (1966): 17-25; and many others.

²¹ See e.g. Chalmers (1996), Jackson (1986); Foster (1991); Gertler (1999, 2006). Also see Kripke, <http://goinside.com/2001/02/25/saul-kripke-genius-logician/>, accessed on October 18, 2012. Cf. Rosenberg (2004).

²² See e.g. Lycan (2009).

²³ See e.g. Gertler (1996, 2006), Lycan (2009), and Arvan (unpublished manuscript).

²⁴ Chalmers (1996).

²⁵ See Russell (1921), Rosenberg (2004), and Arvan (unpublished manuscript).

things interact. Electrons, for example, are defined by what they *do* (they are things that relate to *other* things in very specific ways, viz. to *be* an electron is to orbit atomic nuclei in a particular way). That just seems to be what physical things are. Physical objects and properties are things *that can be seen, touched, smelled, etc.* – things that we and other things *relate* to in particular ways. Qualitative (or “phenomenal”) properties of consciousness, on the other hand, seem utterly *simple*, and therefore *intrinsic*. It seems impossible to fully describe what red looks like – not because red is too complex to describe, but because it is too *simple*. It has aspects that cannot be described *at all* (red simply looks *red* – and what is red? One can only “point” at it: “It looks like *that*”). But now if certain aspects of red cannot be described at all, then they certainly cannot be described by *science* – for science describes the physical world. Thus, physicalism seems false. Consciousness has fundamentally intrinsic, *nonphysical* elements.

Now, it is not my aim here to disprove physicalism. The only relevant question for our inquiry is whether mind-body dualism is a serious epistemic possibility – and clearly, only the most bull-headed physicalist can deny that it is.

§1.5 Evidence for Subjectivity About the Passage of Time

We have already seen that science seems to have trouble making sense of the passage of time. From a physical perspective, all times appear to be equally “real.” Yet, we don’t experience them all as equally real. We experience the past as gone and the future as yet to come. What should we make of this experience – the experience of the ever-ephemeral, forward marching present? A number of philosophers have attempted to argue from our subjective experience of time’s passing to *objective* ontological conclusions about the nature of time. Craig and Schlesinger, for example, argue for

presentism – the view that only the present exists – on the basis of our experience of time’s passing.²⁶ Others, such as Broad and Smith, have defended the “moving-spotlight” theory of time – the view that all times eternally exist, but that the present is marked out by a special, ontologically basic “spotlight” – on a similar basis.²⁷ Finally, Maudlin gives a similar argument for time’s “forward directionality.”²⁸

These sorts of arguments, however, seem clearly predicated upon a mistake. As, L.A. Paul argues, one cannot validly infer ontological claims about the *objective* nature of time from *subjective* experience – for, as Paul points out, even if eternalism were true (if all times objectively exist timelessly), we would *still* have subjective experiences as of time passing.²⁹ There is, however, a compelling way to resist Paul’s argument that Paul does not discuss – and that is by holding that *existing arguments for mind-body dualism* strongly support a hybrid, dualistic theory of time: namely, that (i) eternalism is the correct account of objective, physical time, and (ii) presentism the correct account of time’s *passing*. Let me explain how arguments for mind-body dualism appear to have these implications, and why this dualist view actually is arguably more consistent with our total evidence than any monist theory of time (i.e. presentism or eternalism).

Paul writes, “for reasons of parsimony, we should not postulate the existence of fundamental properties of newness or passage until we have better ontological and empirical reasons to do so.”³⁰ The mind-body dualist, however, *has* such reasons.

²⁶ See William Lane Craig, *Time and the Metaphysics of Relativity* (Philosophical Studies Series 84, 2001); G.N. Schlesinger, *Aspects of Time* (Indianapolis: Hackett, 1980); and G.N. Schlesinger, “How Time Flies”, *Mind*, 91 (1982): 501-23.

²⁷ See C.D. Broad, “Ostensible Temporality,” [1938], reprinted in M. Loux, *Metaphysics: Contemporary Readings* (Routledge, 2001): 272-278; and Quentin Smith, *Language and Time* (New York: Oxford University Press, 1993).

²⁸ Tim Maudlin, *The Metaphysics Within Physics* (Oxford: Oxford University Press, 2007): 135, 142.

²⁹ See L.A. Paul, “Temporal Experience”, *Journal of Philosophy*, 107(7), (2010): 333-359.

³⁰ *Ibid*: 6.

According to the dualist, *consciousness itself* – e.g. what red looks like, what trains sound like, what pain feels like, etc. – is in some sense non-physical. Since, as we have already seen (in §1.1.), science is notorious for having just as much trouble accounting for the seemingly simple, intrinsic appearance “the present” as it has accounting for the seemingly simple, intrinsic nature of *all* conscious phenomenal experiences (e.g. what red looks like), *anyone who finds arguments for mind-body dualism convincing should take those very same arguments to support subjectivism about the passing of time.* After all, the passing of time – the “nowness” of the present – is no more *describable* than what red looks like. Just like the qualia of redness, the qualia of “nowness” seems utterly simple, inexplicable, and so, outside the realm of scientific description or explanation. Accordingly, I submit, *if one is a mind-body dualist, one should also be a time dualist.* One should accept eternalism as the correct account of “physical” time, but presentism as the correct account of “subjective” (or conscious) time. Finally, notice that together, these two views seem to amount to the moving-spotlight theory. For the moving spotlight theory simply says: (a) there is a sense in which all times exist, but also (b) the present is a special “spotlight.” The mind-body dualist has an excellent account of this. *Consciousness* is the moving spotlight. It is a moving present that “runs across” the eternal, timelessly existing multiverse.

§1.6. Evidence for the Further Fact Theory of Personal Identity

The “further fact” theory of personal identity – the view that personal identity is a brute, irreducible fact – does not enjoy a great deal of popularity in philosophy today. However, there are three reasons to take it very seriously. First, it has been carefully defended by some well-respected contemporary metaphysicians (e.g. Chisholm,

Merricks, and Ninan) and philosophers of religion (e.g., Lowe and Swinburne). Second, it has also been defended by a number of “greats” in the history of philosophy, including Butler and Reid, as well as, in my interpretation, by both Kant and Wittgenstein (who both never discuss personal identity *per se* but explicitly accept a “simple”, transcendental view of us as bare subjects of experience). Third, and most importantly,

(A) A strong case can still be made in favor of the further fact view, and

(B) The most common argument against it – its claimed incompatibility with physicalism³¹ – should be rejected by any serious mind-body dualist.

Let me now briefly give the positive case for the further fact view.

The further fact theory is perhaps best introduced by comparison to a more popular theory of personal identity: psychological-continuity theory.³² Psychological continuity theory holds that I persist over time if and only if someone *psychologically continuous* to me persists over time. Here is a simple, science-fiction case to illustrate. James T. Kirk steps on a teleporter on the starship Enterprise. The teleporter disintegrates his body and materializes an exact replica of Kirk’s body on the planet far

³¹ See Ninan (2009) for a summary of and response to this objection.

³² Advocates include Mark Johnston, ‘Human Beings’, *Journal of Philosophy* 84 (1987): 59–83; B. Garrett, *Personal Identity and Self-Consciousness*. (London: Routledge, 1998); Hud Hudson, *A Materialist Metaphysics of the Human Person* (Cornell University Press, 2001); David Kellogg Lewis, ‘Survival and Identity’, [1976], reprinted in his *Philosophical Papers* vol. I (Oxford University Press, 1983); Thomas Nagel, *The View from Nowhere* (Oxford University Press, 1986): 40; Harold Noonan, *Personal Identity*, Second Edition (London: Routledge, 2003); Robert Nozick, *Philosophical Explanations* (Harvard University Press, 1982), Derek Parfit, ‘Personal Identity’, *Philosophical Review* 80 (1971): 3–27; Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984): 207; John Perry, “Can the Self Divide?”, *Journal of Philosophy* 69 (1972): 463–488; Sidney Shoemaker, “Persons and Their Pasts”, *American Philosophical Quarterly* 7 (1970): 269–285; Sidney Shoemaker, “Personal Identity: A Materialist’s Account”, in Shoemaker and Swinburne, *Personal Identity*, (Oxford: Blackwell, 1984); Sidney Shoemaker, “Self and Substance”, in *Philosophical Perspectives* 11, J. Tomberlin (ed.) (1997): 283–319; Sidney Shoemaker, “Self, Body, and Coincidence”, *Proceedings of the Aristotelian Society*, Supplementary Volume 73 (1999): 287–306; Peter Unger, *Identity, Consciousness, and Value* (Oxford University Press, 1990): chapter 5; and Peter Unger, “The Survival of the Sentient”, in *Philosophical Perspectives* 11, J. Tomberlin (ed.) (Malden, MA: Blackwell, 2000).

below. Proponents of psychological continuity don't doubt that it is in fact Kirk. The man on the planet walks like Kirk, talks like Kirk, looks like Kirk, acts like Kirk, claims to be Kirk, has all of Kirk's memories, and is treated as Kirk by everyone.

If only, however, things were so straightforward. The problem with the psychological continuity theory is precisely the fact that it treats perfect psychological *duplicates* of a person as the *same* person. The problem with this idea – and the reason why the further-fact view seems so compelling – is that it seems perfectly conceivable that a perfect psychological duplicate of a person is not the *same* person but rather a *mere duplicate* with an entirely different consciousness. Here is why: each of us seems to experience ourselves not as a set of personality traits but instead as a *bare point-of-view* – as a “vanishing” subject of experience. This was Kant's and Wittgenstein's common point. As Melchert nicely summarizes Wittgenstein's view:

[Wittgenstein] suggests that if you wrote a book called *The World as I Found it*, there is one thing that would not be mentioned in it: *you*. It would include all of the facts you found, including all the facts about your body. And it would include psychological facts about yourself as well: your character, personality, dispositions, and so on. But you – the subject, the one to whom all this appears, the one who *finds* all these facts – would not be found.³³

Or, as Wittgenstein put it in his own words, “The subject does not belong *to* the world; rather, it is a limit of the world.”³⁴ Kant expressed a similar position– for example, “we cannot even say that this [the “I”/the self] is a concept, but only that it is a *bare*

³³ Norman Melchert, *The Great Conversation: A Historical Introduction to Philosophy*, 6th edition (Oxford University Press, 2011): 611.

³⁴ Wittgenstein (1922): 5.632.

consciousness which accompanies all concepts.”³⁵ For Kant, each of us is ultimately a *noumenal*, unknowable “thing in itself.”³⁶ The self is, “merely that unknown X to whom the world appears and by which it is structured into objects.”³⁷ This is why it seems perfectly conceivable that someone else – some other subject of conscious experience – could wake up tomorrow in my bed with all of my personality traits, memories, beliefs, etc., *and still not be me*. I seem to be a *simple, brute subject* of experience, distinct from all particular psychological or physical facts. A perfect duplicate of me could fail to be me precisely because we can imagine “my consciousness”³⁸ flickering off and replaced by a *duplicate* consciousness: some *other* subject having all of my physical and psychological properties. This is, at least, what many great minds (e.g. Kant, Wittgenstein, Butler, Reid, Chisholm, etc.) have thought.

Additionally, the further fact theory of personal identity seems to fit much better with some of our intuitions than rival, physicalist theories of personal identity. For example, animalism (the view that we are our bodies, or perhaps our brains) doesn’t fit well with Locke’s famous “prince and cobbler” example – the case of a “mind-swap” between bodies. This example seems to demonstrate that it is conceivable for minds to swap bodies, which is permitted by the further fact view theory but prohibited by animalism. Insofar as many people have this intuition, the further fact view has serious metaphysical implications uniquely in its favor.

³⁵ Kant (1781): 331.

³⁶ Ibid., as well as Kant (1785): part III.

³⁷ Melchert (2011): 447.

³⁸ These scare-quotes are intended to indicate that, on the further fact theory, “me” and “my consciousness” are *not* two separate things where a subject, “me”, possesses a second object, a consciousness. They are one identical thing: each person is identical to a brute, persisting consciousness.

Clearly, we cannot settle these debates here.³⁹ It is possible that the further fact theory of consciousness may be eventually disproven (though I doubt it). The relevant point for our purposes, however, is simply that the further-fact view – as unpopular as it may be today – remains a serious, epistemically live philosophical hypothesis with some real evidence in its favor.

§1.7. Evidence for a Single, Commonly-Experienced (or “Actualized”) Timeline

The idea that our physical universe is in some important sense “actualized” (in a way that alternative physical universes are not) is clearly an epistemically-live hypothesis, with substantial evidence in its favor. First, it is the default position in philosophy and cosmology, one favored by Occam’s Razor (the principle of parsimony). Insofar as our universe is the only one that any of us have ever actually perceived, Occam’s Razor requires assuming this hypothesis unless and until it is necessary to invoke alternative physical universes (with conscious observers) to explain our evidence. Second, the hypothesis that only one universe (i.e. ours) is “actualized” for conscious observers is suggested by the conjunction of two (epistemically live) hypotheses that inform this paper: the multiverse hypothesis and mind-body dualism. The multiverse hypothesis holds that a vast array of alternative physical universes is necessary to explain quantum physics or eternal inflation. Mind-body dualism, on the other hand, states that consciousness is a nonphysical entity or set of nonphysical properties that exist(s)

³⁹ Some readers might object that the foregoing discussion assumed something that we might not in fact be able to imagine: a *perfect* psychological duplicate of a person. We cannot settle this issue here, but fortunately, it is tangential to our purposes. The further-fact theory does not assume that perfect psychological duplicates are possible. As previously discussed, there appear to be several converging lines of evidence in favor of the further-fact theory (e.g. our seeming distinctness from any particular psychological facts, the latter of which can change incredibly over time while *we* continue to exist as a “bare perspective” on the world). The case of perfect psychological duplicates is only intended as one possible way to *illustrate* the intuitive distinction between persons (as “bare perspectives”) and psychological characteristics.

outside of the physical order. Accordingly, if both of these hypotheses are true, and they are then conjoined with Occam's Razor, it follows that there exists a vast array of physical universes, but only *one* of those universes – our universe – is experienced by conscious observers. This hypothesis is actually quite intuitive, and presupposed in many popular works of science fiction. Let me explain.

Consider a famous case of time-travel and “alternate realities” in science fiction: the film *Back to the Future II*. In this film, the antagonist, “Biff”, takes a sports almanac from the future into the past and gives it to his younger self, who then uses it to win an array of sports bets, accumulating a vast fortune, thereby “changing the future.” The film's main protagonist, “Marty McFly”, then travels back into the past to ensure that the younger Biff never receives the almanac, with the goal of *changing the future back* to its original state. An interesting thing about this example is that a number of different timelines seem to *objectively exist* in some sense, regardless of the characters' actions. When Biff travels back in time to “change the future”, it is implausible that the original timeline – the one that “disappeared” after Biff gave his younger self the sports almanac – is completely erased from existence. The idea that an entire universe is *objectively* erased from existence by Biff's actions, and another whole universe *objectively* brought into existence by his actions, is incredibly implausible for two related reasons. First, if Biff's actions in the past literally *destroyed* a universe and brought into existence an entirely new one, it would be hard to see how Marty's actions in the past could possibly *restore* the original universe. What would original universe be restored *from*? *Nothingness*? How? Secondly, and relatedly, it very idea hard to take seriously the idea that entire universes – each comprising *billions* of years, and vast amounts of matter

and energy – could suddenly “pop” in and out of existence due to a single action by a single person at a single point in time. A far more plausible hypothesis – one that does not require the instantaneous disappearance and appearance of entire universes – is that there a number of objectively existing alternative *physical* universes, and the characters’ actions merely change which physical universe *conscious observers* experience as “actual.”

Now, the *Back to the Future* films are of course fictional, yet they clearly demonstrate how – if mind-body dualism and the multiverse hypothesis are true – a vast array of physical universes may objectively exist while only one such universe (*ours*) is consciously experienced by anyone. Accordingly, since mind-body dualism and the multiverse hypothesis are epistemically-live hypotheses, so too is the hypothesis that our universe is the only universe experienced (or “actualized”) by conscious observers.

§2. The Initial Case for Libertarian Compatibilism

The case for Libertarian Compatibilism then is as follows:

- (A) If Hypotheses 1-3 are correct, then the physical universe is a timelessly existing array of information comprising a vast, possibly infinite number of possible pasts, presents, and futures (i.e. different physical “time lines”).⁴⁰

⁴⁰ Some readers might question whether a truly *infinite* set of distinct physical things (or “timelines”) could exist. What, after all, is the vast “hologram multiverse” supposed to consist in – a vast or infinite array of *information*? These are important questions to ask, but not ones that we can or should attempt to resolve here. The question of whether “actual infinities” are metaphysically possible has been around for millenia (see e.g. Plato’s *Timeaus* and Aristotle’s *Physics*). We cannot enter this debate here, but let us say two things. First, insofar as a number of serious mathematicians and metaphysicians have defended the possibility of actual infinities, their possibility is, at the very least, a live epistemic hypothesis. See e.g. see Leibniz’s *Monadology*, and more recently, Jose Benardete, *Infinity: An Essay in Metaphysics* (Oxford: Oxford University Press, 1964). Second, it is unclear whether the many-worlds interpretation of quantum mechanics requires an infinite, or merely immense but finite, number of worlds.

(B) If Hypotheses 4-6 are correct, then each person's consciousness is a brute, non-physical entity or set of nonphysical properties through which each person experiences time as moving forward.

(C) Thus (from A&B), if Hypotheses 1-6 are correct, then

- a. The physical universe is a timelessly existing array of information comprising a vast, possibly infinite number of possible pasts, presents, and futures (i.e. different physical "time lines"), and
- b. Each person's consciousness is a brute, non-physical entity or set of properties through which one experiences time as moving forward.

Next, we add,

(D) *The Subjective Experience of Free Choice*: each person experiences themselves, from a first-person point-of-view, as though they are making fully free choices – choices not fully constrained or enacted by physical laws.⁴¹

And,

⁴¹ Two notes: First, our choices are clearly at least partially constrained by physical laws (example: I cannot choose to jump to the moon). Libertarian Compatibilism simply denies that our choices are *fully* constrained or enacted by physical laws. Second, different theories of causation understand its nature very differently. Causation is most commonly thought of as enacting (or lawfully necessitating) future events. See e.g. G.E.M. Anscombe, *Causality and Determination* (Cambridge: Cambridge University Press, 1971); Michael Tooley, *Causation: A Realist Approach* (Oxford: Oxford University Press, 1987); Alfred North Whitehead, Alfred North Whitehead, *Process and Reality* (corrected edition), Donald Ray Griffin and Donald W. Sherburne (eds.) (New York: Free Press, 1929); and Zeno Vendler, "Effects, Results, and Consequences", *Analytical Philosophy*, Ronald J. Butler (ed.) (Oxford: Oxford University Press, 1962): 1-15. However, a few theorists have argued that causation is merely a kind of metaphysical constraint on future events. (See e.g. Rosenberg 2004: chapter 9, as well as counterfactual theories of causation, such as David Kellogg Lewis, "Causation", *Journal of Philosophy*, 70 (1973): 556-67; and David Kellogg Lewis, "Causation as Influence", *Journal of Philosophy*, 97 (2000): 182-97.

Forthcoming in *The Philosophical Forum*

(E) If mind-body dualism (Hypothesis 4) is true, then one's subjective experiences *as though* one is makingfully free choices – choices not fully constrained or enacted by physical laws – may be *veridical*.

And,

(F) If the further-fact theory of personal identity (Hypothesis 6) is true, then one could persist over time by virtue of one's consciousness "hopping" from physical body to physical body down a chosen path in the multiverse, on the basis of one's conscious choices.

And,

(G) If hypothesis 7 is correct (Single, Commonly-Experienced [or "Actualized"] Timeline), then all conscious observers experience the same, single "actualized" reality.

So, finally, we have:

(H) Thus (from premises C-G), if Hypotheses 1-7 are true and we have the subjective experience as of making free choices, then Libertarian Compatibilism may be veridical:

- a. The physical universe is a timelessly existing array of information comprising a vast, possibly infinite number of possible pasts, presents, and futures (i.e. different physical "time lines"), and
- b. Each person's consciousness is a brute, non-physical entity or set of properties through which one experiences time as moving forward, and

- c. Free will is each person's consciousness freely choosing *ex nihilo*, in a manner not fully constrained or enacted by the physical laws in any given "timeline" of the multiverse, which physical timeline is "actualized" for conscious observers (such that the joint choices of multiple conscious observers "collapse" many possible futures at any given instant down to a single, commonly perceived "present").

Libertarian Compatibilism is clearly a radical theory. However, as we will now see, it is not mere fancy. It corresponds nearly perfectly to the only working simulation of quantum mechanics that human beings have ever constructed: *massive-multiplayer-online-roleplaying videogames* (or MMORPGs). Furthermore, as we will soon see, Libertarian Compatibilism actually explains *the very existence* of numerous philosophical problems – problems in the philosophy of mind, time, personal identity, and of course, free will. I will argue, as such, that as fantastic and metaphysically extravagant as Libertarian Compatibilism may initially seem, it may well be true. It explains a number of features of our reality that no other physical or metaphysical theory does.

§3. Problems with Existing Interpretations of Quantum Mechanics: A Brief Overview

Libertarian Compatibilism can clearly be understood as a (metaphysical) model of quantum mechanics. First, the multiverse hypothesis it assumes can be understood as modeling the *objective* basis of the quantum wave-function (i.e. the probability space that quantum-physical phenomena comprise). On the other hand, the role it affords to consciousness – its holding that our consciousnesses (as nonphysical entities or sets of

nonphysical properties) jointly winnow down, at every instant, through our perceptions and free choices, the possible “timelines” we can experience, to a single, commonly perceived reality – can be understood as modeling the *quantum wave-function “collapse”* to a single, intersubjectively perceived external world.

As such, Libertarian Compatibilism *merges* two existing models of quantum-mechanics in an intriguing way: (A) the Everett/many-worlds interpretation, which holds that a vast array of alternative universes are all there is to quantum reality (i.e. so there is no actual “collapse” of the wave-function; each world is as actualized as any other); and (B) the many-minds interpretation, a fringe interpretation of quantum mechanics which holds that quantum-reality is nothing more than an infinity of non-physical minds interacting so as to give rise to a common experience of “reality.”⁴² Libertarian Compatibilism takes the most plausible aspect of the Everett interpretation – its assertion of numerous physical universes, to account for quantum indeterminacy – while denying its most implausible feature: its implication that all universes in the multiverse are equally “actualized.” Libertarian Compatibilism accomplishes this by assigning to *consciousness* the role of “universe actualizer” – which is intuitively the most plausible aspect of the many-minds interpretation – while rejecting the many-minds interpretation’s most implausible aspect: its assertion that the universe *itself* consists in nothing more than an infinite number of minds (somehow) joined together in a quantum state. Finally, as we will soon see, Libertarian Compatibilism’s particular way of merging these two interpretations of quantum mechanics – its claim that our reality *emerges* from nonphysical conscious minds interacting with an eternally-existing

⁴² See Dieter H. Zeh, “On the Interpretation of Measurement in Quantum Theory”, *Foundations of Physics*, Vol. 1, Issue 1 (1970): 69-76; and Dieter H. Zeh, “The Problem of Conscious Observation in Quantum Mechanical Description”, *Foundations of Physics Letters*, 13 (2000) 221-233.

physical multiverse, in such a way that the choices of the minds determine which universe is “actualized” for conscious observers – corresponds almost perfectly to the *only working simulation of quantum effects and “reality” that human beings have ever constructed that successfully replicate these phenomena.*

In order to make the case that Libertarian Compatibilism is a better metaphysical model of quantum mechanics than existing interpretations, we must look carefully at a few prominent interpretations. We will focus on three historically (and presently) influential interpretations: the Copenhagen, Everett, and Bohmian models.

§3.1. Problems with the Copenhagen Interpretation of Quantum Mechanics⁴³

The Copenhagen interpretation of quantum mechanics denies that quantum mechanics can provide a true description of an *objective* physical reality. Instead, the Copenhagen interpretation asserts that physical reality is partially *constituted* by (subjective) processes of measuring or observing phenomena.⁴⁴ On the Copenhagen interpretation, physical reality exists in a probabilistic “blur” (or superposition) of possible states until observed, at which point the process of observation collapses the probabilistic blur to a determinate set of values.⁴⁵ A famous example – “Schrodinger’s cat” – can be used to illustrate these ideas more concretely. In Schrodinger’s cat example, we are to imagine:

- (a) Placing a live cat in a completely sealed box, along with,
- (b) A single atom of a radioactive substance, which has some quantum-probability P of decaying in the next hour, and finally

⁴³ For a nice history of the debate over the Copenhagen interpretation, including canonical criticisms, see Stephano Osnaghi, Fabio Frietas, and OlivalFreire, Jr., “The Origin of the Everettian Heresy”, *Studies in History and Philosophy of Modern Physics* (2008): doi:10.1016/j.shpsb.2008.10.002.

⁴⁴ For an introduction, see Hermann Wimmel, *Quantum Physics and Observed Reality: A Critical Interpretation of Quantum Mechanics* (World Scientific, 1992).

⁴⁵ Ibid.

- (c) A Geiger counter that will open a small packet of hydrochloric acid, killing the cat, *if the counter detects the decay of the radioactive atom.*

The Copenhagen interpretation asserts that the cat in this example exists in superposition of life and death – that it is in some sense both alive and dead – until we open the box and take a measurement, at which point our observing the system will cause the quantum-state to collapse into a single determinate set of values (i.e. we will observe that the cat is alive or we will observe that the cat is dead).

We can now see why the Copenhagen interpretation has a number of metaphysically fantastic, notoriously obscure, possibly even incoherent, features.⁴⁶ First, the Copenhagen idea that Schrodinger’s cat is simultaneously alive and dead in the box – in a “superposition”, until observed – seems incoherent. It is logically impossible for something to have a property and not have it, at least in the same spatial location at the same time. Philosophically, the only coherent way to make sense of an object “having contradictory properties” is to posit *two distinct objects*: in this case, a dead cat and a live cat in two different “places” or “phase-spaces” of reality. However, this is essentially the solution asserted by the Copenhagen interpretation’s main traditional rival – the Everett/many-worlds interpretation, which holds that a dead cat and a live cat both exist in *different branches* of the multiverse. Indeed, this issue has been one of the primary motivators for the Everett/many worlds interpretation over the Copenhagen interpretation.

Furthermore, the Copenhagen interpretation has been argued to be incoherent in second way. Once again, the interpretation asserts that a quantum system exists in a

⁴⁶ See e.g., C. Chevalley, “Why do we find Bohr obscure?” In D. Greenberger, W.L. Reiter, & A. Zeilinger (eds.), *Epistemological and experimental perspectives on quantum mechanics* (Dordrecht: Springer, 1999): 598–600.

superposition until measured or observed. Therefore, in order to be viable, the Copenhagen interpretation requires a coherent definition of “measurement” or “observation.” Yet this seems impossible. The most notorious issue with the Copenhagen interpretation is the so-called “Measurement Problem.” The Copenhagen interpretation understands measurement in terms of classical physical processes (i.e. where each particle is understood as having a determinate position and velocity), while at the very same time affirming quantum-mechanics – a set of non-classical processes – as the correct model of physical reality.⁴⁷ This seems incoherent: either the world is fundamentally classical *or* non-classical, not both – unless of course the Copenhagen interpretation is somehow understood as a fundamentally *dualistic* account of reality (as holding that subjective processes of measurement are somehow classical but quantum-mechanical processes non-classical). Accordingly, the Copenhagen interpretation is seriously problematic in a second sense: it either assumes an incoherent conception of measurement, or else a dualist metaphysics, one that treats “measurement” as a fundamentally different kind of thing than the quantum world, without any explanation.

Third, the Copenhagen Interpretation lacks explanatory power, and seems to require hidden, unobservable properties and processes. For even if it were possible to coherently define the notion of measurement, the Copenhagen interpretation still (i) leaves the nature of the wave-function a complete physical (and metaphysical) mystery, in a way that (ii) appears to require properties and processes that are unobservable in principle. As noted physicist Steven Weinberg observes,

⁴⁷ See Osnaghi *et al.* (2008): §2.3 for a review of these problems. Also see Claus Kiefer, “On the Interpretation of Quantum Theory – from Copenhagen to the Present Day”, arXiv:quant-ph/0210152v1, (2002).

The Copenhagen interpretation describes what happens when an observer makes a measurement, but the observer and the act of measurement are themselves treated classically. This is surely wrong: Physicists and their apparatus must be governed by the same quantum mechanical rules that govern everything else in the universe. But these rules are expressed in terms of a wave function (or, more precisely, a state vector) that evolves in a perfectly deterministic way. *So where do the probabilistic rules of the Copenhagen interpretation come from?*⁴⁸

The Copenhagen interpretation only purports to explain the *appearance* of the wave-function collapse to observers, not any objective collapse. When it was first developed, the Copenhagen interpretation treated the appearance of the wave-function collapse as a fundamental, *completely unexplained* process. More recently, a detailed model of the appearance of the wave-function collapse – the model of “decoherence” – has been developed to mathematically represent processes of measurement which can lead observers to measure the many phases (or physical possibilities) in the wave-function as appearing to approach a classical limit (i.e. a collapse down to one determinate actual state).⁴⁹ Decoherence does not, however, account for any *actual* (i.e. objective, or measurement-independent) collapse of the wave-function. On the contrary, decoherence implies that the entire quantum wave-function – its “superposition” in many possible states at once – still literally *exists*, in some metaphysically obscure and unexplained way, even *after* a measurement is taken, *in a manner beyond all possible*

⁴⁸ Steven Weinberg, Steven, “Einstein’s Mistakes”, *Physics Today* (November, 2005): 31; my italics.

⁴⁹ For an extended discussion of quantum decoherence, Raphael Bousso and Leonard Susskind, “The Multiverse Interpretation of Quantum Mechanics”, arXiv:1105.3796v3 [hep-th] 22 (2011).

measurement.⁵⁰ All of this, however, seems hopelessly mysterious. The Copenhagen interpretation appears to provide no (coherent) metaphysical account of the sense in which the wave-function exists in a superposition when not observed. Thus, despite its attempts to remain empirically grounded, the Copenhagen interpretation (a) fails to provide any explanation of the quantum wave-function itself (what the wave-function is independently of our measurements), and thus, (b) cannot avoid positing the existence of metaphysically “hidden”, completely unobservable objects and properties. It is, in short, ontologically extravagant, despite its aims to the contrary.

Finally, the Copenhagen interpretation seems metaphysically extravagant in terms of how it understands ordinary physical objects and properties. Once again, the interpretation holds that our observing the world “collapses” the quantum wave-function to a determinate set of observed values. In Schrodinger’s cat example, our opening the box collapses the “live/dead cat” superposition to one of two realities: a live cat, or else a dead cat. Yet it seems outrageous to make the physical world itself measurement-dependent (“reality is in the observations, not in the electron”).⁵¹ As Einstein once scoffed, regarding the Copenhagen interpretation, “Do you really think the moon isn't there if you aren't looking at it?”⁵² The idea that observations constitute reality has been widely criticized at least since Berkeley made similar philosophical claims in *A Treatise Concerning the Principles of Human Knowledge*.⁵³

In summary, the Copenhagen interpretation is:

⁵⁰ Hew Price, *Time's Arrow and Archimedes' Point: New Directions for the Physics of Time* (Oxford University Press, 1997): 226.

⁵¹ Werner Heisenberg, *Physics and Philosophy: The Revolution of Modern Science* (New York: Harper & Row, 1958): 139.

⁵² A. Pais, “Einstein and the Quantum Theory”, *Reviews of Modern Physics* 51 (1979): 863-914: 907.

⁵³ George Berkeley, *A Treatise Concerning the Principles of Human Knowledge*, [1781] reprinted in *The Works of George Berkeley, Bishop of Cloyne*. A.A. Luce and T.E. Jessop (eds.), (London: Thomas Nelson and Sons, 1957).

(A) *Possibly incoherent* in two distinct ways.

- a. By lacking a coherent operational definition of “measurement” (a fundamental concept in the interpretation), and
- b. By asserting “superpositions” of logically contradictory states (e.g. Schrodinger’s cat is *dead and alive* in the box until observed).

(B) *Lacking in explanatory power*: by failing to provide any explanation of the quantum wave-function independently of our measurements;

(C) *Ontologically profligate*: in that it appears to require positing objects and properties that are unobservable in principle; and finally,

(D) *Metaphysically outrageous*: in that it holds that ordinary physical objects (e.g. the moon, this chair, etc.) lack any determinate properties until observed.

§3.2. Problems with the Everett/multi-worlds interpretation of Quantum Mechanics

The Everett-“many worlds” interpretation of quantum mechanics purports to solve the aforementioned problems with the Copenhagen interpretation by providing a coherent metaphysical explanation of the quantum wave-function and our observations of its apparent “collapse.” The many-worlds interpretation holds that the quantum wave-function *just is* a vast array of physical worlds, with each world corresponding to a state of the wave-function. Using Schrodinger’s cat example, the many-worlds interpretation holds that there are (parallel) worlds in which the cat is alive, dead, etc., for every state of the wave-function. Incoherence is avoided by holding that each cat exists in a different universe – so that there is no single cat that is simultaneously alive and dead. The many-worlds interpretation thus appears to improve upon the Copenhagen interpretation’s first two problems: the many-worlds interpretation is neither

incoherent nor lacking in explanatory power. It explains the wave-function, and our observations of its “collapse”, in terms of the existence of many possible universes and our existence in one of them (thus leading us to perceive the wave-function as “collapsing” down to a determinate set of values – the values of our universe).

The coherence and explanatory power of the Everett interpretation, however, come at a great cost. First, the Everett interpretation is ontologically profligate. It asserts the existence of a vast array of physical universes that none of us have ever observed or measured.⁵⁴ Second, it makes the extravagant assertion that every universe in the multiverse exists in the same metaphysically “actualized” way (in other words: all of the worlds it asserts are equally “real”). Allow me to explain.

The many-worlds interpretation asserts that (a) its many “worlds” are all there is to the quantum wave-function, and (b) that the quantum wave-function is the fundamental description of reality. Accordingly, the many-worlds interpretation implies that every universe is ontologically on a par with every other. It entails that there is a world in which “I” – a fully conscious near-duplicate of me – am President of the United States. It also entails that there is a world in which “I” am an Olympic athlete; a world in which “I” am a serial killer; etc. All of this seems metaphysically profligate. It simply goes far beyond empirical observation. The only universe that any

⁵⁴ Several possible experiments have been proposed for testing the many-worlds interpretation against the Copenhagen interpretation – see e.g. David Deutsch, “Three Experimental Implications of the Everett Interpretation”, in R. Penrose and C.J. Isham (eds.), *Quantum Concepts of Space and Time* (Oxford: The Clarendon Press, 1986): 204-214; Michael Lockwood, *Mind, Brain & the Quantum* (Oxford: Basil Blackwell, 1989), and Lev Vaidman, “On Schizophrenic Experiences of the Neutron or Why We Should Believe in the Many-Worlds Interpretation of Quantum Theory”, *International Studies in the Philosophy of Science*, 12 (1998): 245-261. However, none of these predictions are presently feasible to test experimentally. It is also important to note that these predictions are different than those of Libertarian Compatibilism (see §4.7). Thus, it is possible (at least in principle) to empirically test which interpretation of quantum mechanics corresponds to reality.

of us presently have any evidence of being “actualized” is our own. If there is an alternative metaphysical hypothesis that can explain quantum-physics just as well as the many-worlds interpretation *without* affirming the existence of a vast, possibly infinite array of “equally” actualized worlds and conscious duplicates of each of us, then that alternative interpretation is to be favored on grounds of parsimony (i.e. Occam’s Razor).

§3.3. Problems with the Bohmian Interpretation of Quantum Mechanics

Let us now turn to the De-Broglie-Bohm theory of quantum mechanics, otherwise known as Bohmian Mechanics.⁵⁵ According to Bohmian Mechanics (and other “hidden variable” theories), quantum mechanics must be understood in terms of completely unobservable classical particles determining how the quantum wave-function collapses. The main allure of Bohmian Mechanics is that it aims to coherently explain how all of quantum reality – the wave function and its collapse – could be due to “hidden” *classical mechanics*. Unlike the Copenhagen interpretation, which leaves the nature of the quantum wave-function completely unexplained, Bohmian Mechanics purports that quantum mechanics can be understood in terms of actual, *completely deterministic* processes. Additionally, it seems not to violate ontological parsimony as much as the many-worlds interpretation, which posits a vast array of unobservable universes.

⁵⁵ See David Bohm, “A suggested Interpretation of the Quantum Theory in Terms of Hidden Variables, I and II,” *Physical Review* 85 (1952). Cf. Albert Einstein, B. Podolsky, and N. Rosen, “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?,” *Physical Review* 47 (10) (1935): 777–780. Hidden variable theories, including Bohmian Mechanics, dramatically fell out of favor for several decades after von Neumann was thought to have proven that all such theories are impossible – see von Neumann John. *Mathematische Grundlagen der Quantenmechanik* (Berlin: Springer, 1932). However, Bell is now credited with having refuted von Neumann’s objections, and Bohmian Mechanics is once again thought by many to be a valid challenge to the more orthodox Copenhagen and many-worlds interpretations – see John Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge: Cambridge University Press, 1988).

Bohmian Mechanics wishes to understand quantum mechanics entirely in terms of classical processes within *this* world.

However, Bohmian Mechanics only accomplishes these things at great cost. First, it (like other hidden variable theories) the existence of particles and processes that are *completely unobservable in principle*. Second, Bohmian mechanics asserts that the particles and processes we observe – quarks, bosons, etc. – interact in some *mysterious and undetectable* way with these hidden variables to cause the quantum wave-function to immediately, *and randomly*, assume one of the many possible values expressed in the equations of quantum mechanics. In short, Bohmian mechanics posits (A) a completely unverifiable kinematics (or processes) beyond *any possible measurement*, and (B) a completely unverifiable and mysterious mechanism through which our taking measurements causes those “behind-the-scenes” kinematics to “collapse” to a single set of observable values, post-measurement. Finally, two independent observers measuring the same quantum system at “the same time” (subject to Einsteinian relativity) can measure *different* “post-collapse” values, making it all the more mysterious how there can possibly be the kind of single, objective, determinate “behind-the-scenes” kinematics Bohmian Mechanics asserts (if there were such a set of single, deterministic processes, one would expect different observers under precisely the same conditions to observe identical post-collapse quantum values – but this isn’t empirically the case).⁵⁶

§3.4. Summary: Problems with Traditional Interpretations of Quantum Mechanics

Although we cannot discuss every existing interpretation of quantum mechanics, we have just seen that three prominent, traditional interpretations – the Copenhagen

⁵⁶ This is a well-known, and empirically verified, feature of quantum physics.

Interpretation, Everett/many-worlds interpretation, and Bohmian Mechanics – all assert the existence of objects and properties that are *unobservable in principle*. In other words, there appears to be no coherent way to understand quantum mechanics *without* positing unobservable objects and properties. This is relevant for our purposes because the most obvious objection to Libertarian Compatibilism is that it is “metaphysically fantastic”, positing a vast multiverse, nonphysical consciousness, and “further facts” about personal identity that “go beyond” our empirical evidence. However, we have just seen that *all* leading interpretations of quantum mechanics posit unobservable entities. The real question, then, is whether the unobservable entities that Libertarian Compatibilism posits are in any way more objectionable than the ones posited by rival interpretations. I will now show that Libertarian Compatibilism in fact has *more* explanatory power, at less metaphysical “cost”, than any rival interpretation of quantum mechanics. Finally, we will see that Libertarian Compatibilism makes distinct, testable predictions that (if verified) would lend it empirical support over rival interpretations.

§4. Libertarian Compatibilism, Quantum Mechanics, and Massive Multiplayer Online Roleplaying Videogames (MMORPGs)

We will now see that Libertarian Compatibilism is no more ontologically extravagant than more traditional interpretations of quantum mechanics, while possessing far more explanatory power. Accordingly, I shall submit that Libertarian Compatibilism is a better model of quantum mechanics than traditional interpretations.

Let us begin by briefly reflecting on the nature of theories: what they are, what they aim to do, and what it is for a theory to successfully explain phenomena. Roughly speaking, theories are *predictive models*. A theory of climate change, for example, will

specify a number of parameters (CO2 levels, sunlight, etc.), how those parameters relate to one another (e.g. rising CO2 will increase cloud cover, reflecting sunlight back towards Earth in a positive feedback mechanism, etc.), and finally, make predictions about what will happen given certain inputs for each parameter (e.g. if CO2 levels = 100 parts-per-million and sunlight remains constant, there will be a mean increase in troposphere temperature of 2 degrees Celcius over ten years). A *successful* theory, then – a veridical one, with explanatory power – is thus a *model that accurately predicts observed phenomena*. For example, if we ran two climate models – C1 and C2 – and our observations over a period of time agreed with the predictions of C1 but not C2, we would conclude that C1 is a more accurate theory (at least relative to the time period observed). Finally, the more things it (accurately) explains, the *greater* a theory's explanatory power. If chemistry accurately predicts how carbon dioxide behaves but quantum-physics accurately predicts deeper physical mechanisms that cause carbon dioxide to behave the way it does, then quantum-physics is a more complete theory of carbon dioxide's behavior than chemistry.

Let us now carefully examine the model of reality proposed by Libertarian Compatibilism. Nick Bostrom has argued, in an ingenious and widely discussed paper, that we may all be living inside of a giant computer simulation – that we may actually be simulated agents analogous to the kinds of simpler simulated “people” we have created within virtual reality role-playing games (such as in the game *The Sims*).⁵⁷ Similarly, David Chalmers has argued that “The Matrix Hypothesis” – the hypothesis that we are all living inside of a giant simulation – is metaphysically identical to a hypothesis about

⁵⁷ Nick Bostrom, “Are You Living in a Computer Simulation?”, *Philosophical Quarterly*, vol 53, No 211 (2003): 243-255.

the nature of our world (i.e. the “Metaphysical Hypothesis”) that many ordinary people already accept. Chalmers’ Metaphysical Hypothesis holds that (1) reality is fundamentally computational, (2) our cognitive systems are distinct from cognitive processes (but somehow interact with them), and (3) our reality was created by a being or beings (e.g. God) outside of our space-time.⁵⁸ Although Bostrom’s and Chalmers’ papers have received a great deal of discussion, I will now argue that neither of them (nor their commentators) have yet gone far enough “down the rabbit-hole.” For as I will now show, actually existing simulations of Libertarian Compatibilism *accurately predict, reproduce, and so explain* (1) every general feature of quantum mechanics, as well as (2) the emergence of several longstanding philosophical problems in the philosophy of mind, personal identity, time – and, of course, free will. Therefore, I shall submit that Libertarian Compatibilism corresponds to the *only working model* human beings have ever created that predicts and replicates “the appearances” of our world.

§4.1. Massive-Multiplayer-Online-Roleplaying-Videogames (MMORPGs) as a Simulation of Libertarian Compatibilism

Let us recall the basic elements of Libertarian Compatibilism. According to this theory, reality consists of:

- An eternally existing array of (physical) information comprising a vast array of possible pasts, presents, and futures.
- External, nonphysical entities – each person’s consciousness – reading that information in “real time”, in such a way that,

⁵⁸ David Chalmers, “The Matrix as Metaphysics”, In Christopher Grau, (ed.), *Philosophers Explore the Matrix* (Oxford University Press, 2005).

Forthcoming in *The Philosophical Forum*

- The joint choices of all of the nonphysical entities – choices which are in no way causally determined by the physical information – interact so as to generate a single, commonly perceived “external reality.”

Let us now discuss the mechanics of MMORPGs. Existing MMORPGs are comprised by a large number of independent game consoles “system-linked” together over the Internet.

- First, each game console contains and plays a game DVD. Each game DVD consists of an array of information containing a vast range of possible “pasts, presents, and futures” within the game (i.e. possible positions and environments players can exist in and perceive, “rocks”, “automobiles”, etc., within the simulation).
- Second, each game player’s experience of the simulated environment “in real time” (i.e., how the events of the game play out) is the product of an entity outside of the information on the DVD – namely, the laser-apparatus and computer processor of the game console. This laser-apparatus reads a particular string of physical information on the DVD, and the computer processor thereby “actualizes” it. From the perspective of anyone in the “physical” environment within the simulation, these mechanisms would count as nonphysical.
- Finally, each player’s choices on their own game console are processed by the network so that the joint choices of each player – choices not causally determined by any “physical law” within the simulation – lead each game console hooked up to the system to read complementary lines of code on the DVD to ensure that the physical laws of the simulation are not broken. This

allows each player on the network to experience the same virtual environment from different points of view.

The point here is that these MMORPGs embody assumptions that are *almost perfectly analogous* to Libertarian Compatibilism's model of our world. Libertarian Compatibilism assumes a "multiverse" of information "written" on a cosmological horizon. An MMORPG game DVD is *just that*: an array of information comprising a number of possible pasts, presents, and futures. The "multiverse" assumed by Libertarian Compatibilism and an ordinary MMORPG's game DVD are, then, *the same kind of thing*: series' of information comprising "possible pasts, presents, and futures" (MMORPG game DVDs are merely *smaller, containing far less information*, than Libertarian Compatibilism's multiverse). Now consider Libertarian Compatibilism's assumption that consciousness is a nonphysical medium which reads physical information off to each person's subjective awareness. The laser-apparatus and computer processor in a MMORPG game console are *strongly analogous* to this feature of Libertarian Compatibilism. Observers within an MMORPG game would have no knowledge of the laser apparatus and processor outside of their "physical" world that, *in actuality*, comprise their experience of their "physical" world. While these mechanisms would give rise to (and indeed, comprise) their experience of their simulated world, the mechanisms themselves are not contained within the information that comprises their "physical" reality (i.e. the physics of their observed world). Finally, MMORPGs correspond to the Libertarian Compatibilist model of free will and our experiences of "quantum collapse." When I play a game of Halo (a famous MMORPG), the free choices made by myself and other users outside of the simulation (i.e. the video

game) *cause* the events within the simulation to unfold as they do (through processes that would be “invisible” to every observer within the simulation). Further, all of the features of the simulated environment (blowing leaves, rain, etc.) are represented *slightly differently* by each game system in a system link. Until the next instant of the game occurs (as a result of each player’s new choices), all game-systems become nearly aligned (due to each individual game system processing information sent to it from all of the others). The “physical” properties of those objects in the simulation (e.g. the exact placement of a single leaf, etc.) will (a) appear to have *many possible values* to each different player, until (b) each game system “decides” (again, in virtue of its system link to the others, and users’ choices outside the simulation) which precise value it will process at the next instant⁵⁹ (e.g. fixing the placement of the leaf determinately).⁶⁰

⁵⁹ And what is an “instant” in Libertarian Compatibilism, and in MMORPGs? Answer: the smallest “bit” of information – or quanta – that comprise “physical information”. For more, see the notion of *decoherence* in quantum physics, where each “splitting” of the multiverse into different branches at every instant is given a quantum-mechanical description. See W.H. Zurek, “Decoherence, einselection, and the quantum origins of the classical,” *Review of Modern Physics*, 75 (2003): 715-775), M. Schlosshauer, “Decoherence, the Measurement Problem, and Interpretations of Quantum Mechanics,” *Review of Modern Physics*, 76 (2004): 1267-1305), and especially Bousso and Susskind (2011).

⁶⁰ It was brought to my attention by an anonymous commenter that (A) I am describing a *peer-to-peer* (P2P) network connection, in which clients communicate between directly between one another and in which there is no dedicated server that establishes an authoritative set of information, (B) many online games today (e.g. World of Warcraft, Command & Conquer) *do* use this format, but also (C) many games today (including Halo 3) utilize a *different* format. Halo, in particular, is coded to depend on a *dedicated server*, in which there *is* one set of authoritative information (i.e. one objective set of information that different consoles on the system are responsive to and aim to represent in tandem – though there are often divergences, or information “lag”, situations in which different observers witness slightly different values and events). Fortunately, as this reader concedes, none of this is a strike against Libertarian Compatibilism. Dedicated servers are utilized for contingent reasons of processing ease and economy. Very large-scale MMORPGs – the kind that contain thousands of users across the world, and which are most akin to our reality – *still utilize the P2P format*, the format that corresponds most directly with Libertarian Compatibilism. Because Halo is one of the few MMORPGs that I am directly familiar with, I will continue using it in my discussion. Readers should simply note that although Halo *in fact* uses a dedicated server, we can discuss it *as though* it uses a P2P connection, as many similar games in fact do. For the sake of clarity and simplicity, this will be my approach here. For helpful introductory discussion of these issues, see

Notice that all of the events that transpire within MMORPGs must appear to observers within the simulation to be exist within a “causally closed” order and be determined by other physical events, thus making genuine (libertarian) free will appear to be impossible to any observer inside the simulation. This is illustrated by the fact that, once I have finished a game of Halo, I can go into “theater” mode and have the game replay everything that transpired in the game I just completed. Because the players’ choices in an MMPORG cause their game console to read particular *unbroken series* of information – a series which can be *recorded and played from start to finish over and over again* – any observer “trapped” within such a simulation would perceive their world as being a completely unbroken string of “physical” information: a “causally closed” and deterministic series of events where each event would appear to inexorably lead to the next. In other words, the choices we make outside of the Halo simulation show how genuine free will can give rise to the apparent impossibility of free will to observers within the game. As such, existing MMORPGs demonstrate exactly how the Libertarian Compatibilist model reconciles determinism, causal closure, and genuine freedom of will. From the reference-frame of any observer within an MMORPG, free will appears impossible: *every* event in their “world” seems inexorably determined by previous events. However, in our higher frame-of-reference outside of the simulation, our choices – choices undetermined by the “laws” that govern the simulation – are what in fact determine how the events in the simulated world play out.

There is one crucial difference between Libertarian Compatibilism and MMORPGs, one which may give rise to an objection to the parallel I have been drawing

<http://gafferongames.com/networking-for-game-programmers/what-every-programmer-needs-to-know-about-game-networking/> (accessed on 10/29/2012).

between them. Although MMORPGs model Libertarian Compatibilism's basic idea – the idea that the appearance of causal closure and determinism within a “world” can arise from freely-made choices outside of that world – MMORPGs only push back the problem of free will to a higher level. For while our “choices” outside of MMORPGs are free *relative to the physics of the simulation* (i.e. the video game), the very problem of free will (in our world) is that *our* choices appear to be casually determined by the physical laws of *our* world. Let us return to this worry later, after we have examined the theory's implications in more detail.

We have just seen that existing MMORPGs – online role-playing videogames that human beings have already created – correspond in almost every respect to Libertarian Compatibilism's model of our reality. As such, we can understand MMORPGs *as* simulations of Libertarian Compatibilism, and accordingly test Libertarian Compatibilism's empirical predictions for our world. After all, if our world is – as Libertarian Compatibilism says – a kind of *simulation* where free choices in a higher-reference frame give rise to the *apparent* impossibility of free will in our reference, Libertarian Compatibilism may make distinct empirical predictions: predictions that it alone makes, and which, if confirmed, would *verify* the theory. And indeed it does. We can test Libertarian Compatibilism's predictions in roughly the same way that we test climate simulations, which of course we test against observations of Earth's actual climate. First, as we have just seen (and will see further in §§4.2-4.7), MMORPGs reproduce *all* of the essential “appearances” of our world: the quantum uncertainties we observe, a variety of philosophical problems spanning the philosophy of mind, time, etc. As such, Libertarian Compatibilism actually predicts and explains the appearances of

our world *better* than any existing physical or metaphysical theory. Unlike existing theories of quantum mechanics, which take quantum phenomena to be fundamental features of reality, Libertarian Compatibilism actually provides a deeper explanation of where those features come from (viz. quantum features naturally emerge from *any* simulation comprised by a “system link” between independent media – see §4.2). Finally, and very importantly, as we will see in §4.7, Libertarian Compatibilism makes *three* unique predictions in quantum physics, and thus, is amenable to empirical verification and falsification.

§4.2. How Libertarian Compatibilism Provides a More Powerful Explanation of Quantum Indeterminacy Than Existing Interpretations of Quantum Mechanics

For the sake of clarity, let us discuss a particular (and quite famous) MMORPG in existence today: “Halo.” Within Halo’s online multiplayer game, each player adopts the role of a “Spartan” super-soldier. The aim of the game is to move around the virtual environment – one can hide behind rocks, drive vehicles, etc. – and “kill” as many other players as possible. With these basic points in mind, let us investigate the “physics” of the Halo simulation from the perspective of a “Spartan” observer.

A notable aspect of the Halo simulation is that its programmers went to great lengths to replicate various features of our world. Take ballistics, for example. “Halo 3” actually models the *ballistics* of every bullet one shoots from one’s “gun” in real time. It subjects each and every simulated bullet to effects of (simulated) gravity, wind, etc. (each and every “bullet” literally “flies through the air” inside the simulation). Here, then, is an interesting question: when one shoots a bullet in Halo, where does the bullet *really exist* within the simulation? As any player of Halo would attest, the answer to this

question is not simple at all. Each player's *own* game console is playing a distinct copy of the game DVD. Accordingly, when I shoot a "bullet" at another player on my console, *my* console will "code" that bullet as existing at a particular "space-time" location in the simulation. For example, my console will represent my bullet as being at spatial location, *S*, and as having occurred at some time *t* relative to other things in the simulation (e.g. I "shot" after player 2 "shot", etc.). But will other players' game consoles process my "bullet" as being in the exact *same* space-time location within the simulation? The answer is *no*. Because the various game consoles system-linked together are distinct, and have to transmit information to each another over the Internet in order to give every other game console "real time" information on the physical position(s) of objects within the simulation, each game console has to *approximate* the information on all of the others. Accordingly, each player may experience a *slightly different* "space-time" position of the same "bullet."⁶¹ Moreover, because each individual game system has to compute these things probabilistically, predicting the "bullet's location" based on information given to it (the console) by other consoles, each observer within the simulation (e.g. me playing the simulation in my bedroom, you playing the simulation in your bedroom) will inevitably only experience a probabilistic account of where the bullet "really is." This is how actual (peer-to-peer networked) MMORPGs work.

⁶¹ Again, see footnote 59. I have learned that the Halo games utilize a "dedicated server", which serves as the ultimate arbiter of a bullet's "objective" location in the simulation (which all other systems linked to the simulation approximate). Peer-to-peer simulations, on the other hand, have *no* system linked to the simulation which represents the "objective" location of objects, and thus correspond better to my discussion here. I only use Halo 3 as my example because it is one of the few MMORPGs I am familiar with. I am discussing it *as if* it utilizes a peer-to-peer (P2) networking model, which – as I mentioned earlier – many other games (World of Warcraft, etc.) utilize. The fact that Halo 3 utilizes a dedicated server in no way affects the substance of my discussion.

Insofar as this is true (i.e., there is no determinate location of the “bullet,” only a probabilistic one), existing MMORPGs *replicate* the “quantum phenomena” we experience in our world. First, MMORPGs replicate the probabilistic nature of quantum mechanics: an observer within the Halo simulation can only know where a “bullet” is in his/her environment in a probabilistic way – not a determinate one. Second, MMORPGs replicate a baffling measurement problem in quantum mechanics: the fact that different observers of the same physical system can “simultaneously” arrive at different measurements. Again, quantum physics holds that if you and I attempt to observe a quark “simultaneously”, you may measure it as having properties P and I may measure it as having slightly different properties P^* . This is exactly the same case with the “bullet” example within the Halo simulation. Again, my console computes “where the bullet really is” using feedback it receives from other consoles: the “measurements” that each system makes at any given instant is reciprocally affected by the measurements of all other systems attached to the simulation. And again, this is precisely what quantum-mechanics tells us: my observations of a quark will affect your observations of it at that very instant, and vice versa.

MMORPGs thus replicate several key features of quantum mechanics that we perceive in our world. MMORPGs are, therefore, the *only* working model that human beings have ever actually created to predict and replicate our observations of quantum reality. Further, MMORPGs demonstrate that two theories previously discussed in this paper – eternalism and the multiverse hypothesis – are both far better understood in the context of Libertarian Compatibilism than on their own. Since eternalism states that past, present, and future objects all exist timelessly, it is unclear how it can model our

experience of the “passing of time.” Proponents of eternalism have tried to account for the appearance of time’s passage by holding that time will “appear” to pass to any observer at any point in the eternal order, since the eternal order simply is a temporally ordered series of objects and events. However, MMORPGs clearly reveal that this account cannot be correct. We have just seen that MMORPGs model eternalism. From the perspective of any observer within the Halo simulation, the information encoded upon each and every Halo game DVD is an *eternally existing order* of (possible) events. Yet the series of “eternally” existing events are not themselves sufficient to account for any *observations* of time’s passing in the simulation. The temporally-ordered events – the events encoded on the game DVD – *exist* “eternally” from the perspective of observers of the simulation but are only *perceived to be moving forward* in “real time” once the DVD is *played* (i.e. on the game console).

In short, the only working simulations of “eternalism” that we have ever constructed – the series of events encoded on MMORPG DVDs – suggest that traditional forms of eternalism are *false*. MMORPGs show that an “eternally” existing, temporally-ordered series of information needs something additional in order to simulate the experience of the “passing of time”: a *medium* to read those “eternal” events in real time (in this case, the game-console’s laser apparatus and processor). MMORPGs thus suggest that “time” must have two elements: an “eternal” array of information, and a *moving “spotlight”* to account for “the present” and the forward-march of moving time.

MMORPGs pose a similar problem for the traditional multiverse hypothesis (i.e. the multiverse hypothesis considered unattached to the rest of Libertarian Compatibilism). The traditional multiverse hypothesis (i.e. the Everett interpretation of

quantum mechanics) states that every state of the quantum wave-function corresponds to an *actually existing physical universe*, none of which are any more “actualized” than any of the others (i.e. a multiverse in which there are countless *conscious* duplicates of you, I, and everyone else). Yet MMORPGs show that we can account for quantum phenomena *without* affirming that all universes are equally “real” or “actualized.” MMORPGs show that we can account for the appearance of quantum indeterminacy in our world by positing a multiverse in which only one “timeline” is actualized. MMORPGs, our only working model of quantum phenomena, show that we can account for every relevant physical observation without being as metaphysically extravagant as the multiverse hypothesis.

§4.3. How Libertarian Compatibilism Predicts and Explains Actual Philosophical Debates in the Philosophy of Time

We have just seen, through understanding MMORPGs as simulations of Libertarian Compatibilism, that Libertarian Compatibilism provides a powerful explanation of quantum indeterminacy. But this is not all the theory explains: it also explains the very existence of an entire array of philosophical problems. The more observed phenomena a theory explains, the more powerful (and successful) that theory is – so accordingly, Libertarian Compatibilism is far more powerful than any other physical or metaphysical theory in existence. Philosophers are accustomed to investigating philosophical problems. I believe that Libertarian Compatibilism is the first unified explanation of why those problems even exist at all.

Let us begin with the philosophy of time. We have seen that philosophers have debated the nature of time for many years. Do all objects and events eternally exist, or

is there some kind of “ephemeral present” that moves forward in “real time”? Libertarian Compatibilism explains why these questions arise. Consider the perspective of agents within an MMORPG. When my character “walks” around the online world of Halo, there is a “moving present” from his perspective. He moves through the world “in real time.” Accordingly, if we imagine ourselves “trapped” within the world of Halo as a character (having never seen or experienced reality outside of the simulation), we would have perceptions *as of* time passing. And yet...it would be very mysterious how this “passing” could occur without *passing into something* (i.e. the future) – just as it is with time in our world. In other words, people in the world of the simulation would be confronted with *all of the same phenomena about time* that we philosophers debate in our world. If we were observers in an ordinary MMORPG, we would inevitably debate “presentism”, “eternalism”, and the “moving-spotlight” theories of time, just as philosophers do in *our* world. Libertarian Compatibilism thus explains the very existence of these philosophical problems about time in our world. Remember, a good theory of some phenomenon *predicts* that phenomenon. Because MMORPGs *reproduce* the very philosophical problems about time that we experience in our world, and MMORPGs conform to Libertarian Compatibilism, Libertarian Compatibilism explains the very existence of philosophical problems about time in our world (i.e. the philosophical questions we raise about time, the debates that ensue, and the positions defended).

§4.4. How Libertarian Compatibilism Predicts and Explains Actual Philosophical Debates in the Philosophy of Mind

Let us now explore the philosophical problems about the nature of mind and consciousness that would present themselves to agents living in a MMORPG. A player can investigate the nature of the “physical” world in the online Halo simulation, and observe (as scientists in our world do) how gravity operates. One can thus observe create theories of how the Halo world’s “physics” work. Anyone doing this kind of “Halo Science” (as I will call it) would quickly observe that the physics of the Halo world is functional in nature. Halo Scientists would construct laws to explain how things in their reality behave, just as we do. Yet something would seem amiss to them about their place in the world as conscious observers. They would sense, just as we do, that there is an “explanatory gap” between their physical theories and their subjective experience. For again, in an MMORPG, there in fact *is* an epistemological and metaphysical gap between the “physical” world and the experiencing subject. The *physics* of an MMORPG is information encoded upon a DVD. Halo Scientists would be able to investigate and discover the various “physical” relations between these bits (or “quanta”) of information. However, their science would inevitably leave something out. The system that *reads* the DVD (i.e., the laser-apparatus and processor and the game console) – the system that gives rise to perception, and the experienced passage of time in the Halo world – would be *inaccessible* to Halo Science. It would be inaccessible because that very system – the system that comprises the point-of-view of each person in the simulation – is *in fact* an entity outside of the physical information the players experience (the system fundamentally *comprises* their experience of every “physical”

thing in their world). In short, given that all of their experiences are (partly) comprised by a system *outside* of their “physical” world (i.e. the laser-mechanism and processor *reading* the information on the game DVD), Halo Scientists would face the same frustrating philosophical issue that we face: nearly everything in their world would *seem* “physical”, and yet they would have the sneaking feeling that their consciousness, or subjective experience, could not possibly be physical – that there is an “explanatory gap” between Halo Science and their own subjective points-of-view as observers. *And they would be right.* The laser apparatus and processor that give rise to subjective experience and the passing of time in the Halo World *exist and are in fact* “nonphysical” from their frame of reference within the simulation.

Libertarian Compatibilism thus predicts and explains the very existence of the mind-body problem in philosophy. Anyone living in a Libertarian Compatibilist world – i.e., any observer in an MMORPG – would ask the *very same questions* (and have the same debates) about the nature of mind and consciousness that we do. Given that mind-body dualism is true in Libertarian Compatibilism, I submit that our only working model capable of replicating the mind-body problem is one in which dualism is in fact true.

§4.5. How Libertarian Compatibilism Predicts and Explains Actual Philosophical Debates About Personal Identity

Now let us turn to the kinds of questions that agents in an MMORPG would ask about personal identity. First of all, agents in an MMORPG would find some attraction in the “bodily” criterion of personal identity, for they would perceive each other as comprised by certain physical forms (or bodies). To illustrate, when I encounter a player in black

armor in the Halo World, I assume it is my friend who goes by the name “Halo God.” But as an agent in an MMORPG, I might wonder, “What would happen if Halo God stepped in a teleporter? Would the same person emerge on the other end?” Since, in the game of Halo, people frequently step into teleporters and emerge apparently unscathed on the other end, agents might also be naturally drawn to the “psychological” criterion of personal identity. They might ask, “Well, look, if the guy who walks out of the other end of the teleporter acts in just the same way as the guy on the front end – possessing all of the same psychological characteristics – wouldn’t it have to be the same person?” Yet some agents might have the nagging suspicion that personal identity is something more. Each agent has a subjective point-of-view, so they might well wonder (just as we do) whether a psychological duplicate that emerges on the other side of a teleporter is a mere duplicate, not the same person who entered. In other words, they might suspect that personal identity is some kind of brute, further fact – a kind of “sameness of consciousness” (though they might have trouble explaining what the sameness could be). *And they would be right.* There would indeed be “further facts” to personal identity outside of their “physical” reality: namely, the particular game console (or medium) upon which each person’s subjectivity is “running”, and of course, me, the “outside” user. These things are not physical or psychological facts within the simulation, but they intuitively comprise personal identity of the characters in the simulation. My character is the same guy from moment-to-moment not because of anything physical or psychological in his “world,” but because *I* am playing him on a *particular console*, both of which are “further facts” from the perspective of the Halo World.

Libertarian Compatibilism thus explains the kinds of philosophical questions and debates that people in our world have about mind-body dualism. Furthermore, the only actual simulation we have ever constructed of our world – MMORPGs – verifies the “further fact” view.

§4.6. How Libertarian Compatibilism Predicts and Explains Actual Philosophical Debates the Free Will Debate

Finally, Libertarian Compatibilism predicts and explains the free will debate. Halo Scientists, if they were investigating their physical world, would almost certainly reach the conclusion that their reality is either fully deterministic or causally closed. They would see – as we do in our world – that every event in their world appears enacted by every other. They might say, “If one were to rewind our world to the very beginning, every event would have occurred just as it has.” Because every event would appear to have a sufficient cause, they would conclude that there is no “room” for genuine free will in their causal order. Presumably, they would then debate “compatibilism”: whether free will is compatible with a deterministic or causally-closed order. In short, agents in an MMORPG would ask the same questions about free will and face the same problem of free will that we do. Thus, since MMORPGs approximate Libertarian Compatibilism, our only working simulation of the *problem* of free will is Libertarian Compatibilism. Libertarian Compatibilism explains the very existence of the problem it purports to resolve.

§4.7. Libertarian Compatibilism’s Empirical Predictions

One of the most common complaints about existing interpretations of quantum mechanics is that they do not make any testable predictions. How, for example, might

we confirm or falsify the many-worlds interpretation? According to the theory, the vast number of other existing universes that exist are inaccessible to us. We can only observe our own universe, and observers in other universes will only be able to observe theirs. Bohmian Mechanics arguably makes one testable prediction: the pilot-wave hypothesis (though it is unclear whether these predictions can be confirmed, given problems of measurement).⁶² Libertarian Compatibilism, on other hand, makes *three* empirical predictions that no other models of quantum mechanics do.

Libertarian Compatibilism makes the following empirical claims: (1) the world we perceive is a *hologram* (viz. the Holographic Principle), (2) the world we perceive is a *simulation* (viz. the conjunction of all the hypotheses discussed), and (3) we have *libertarian* free will in a higher reference-frame that is nevertheless “masked” by apparent causal-closure in our lower-level reference frame (within the simulation). As a matter of fact, all three predictions can, in principle, be tested. We have already seen that the Holographic Principle makes determinate predictions, and that these predictions are already in the process of being tested.⁶³ Second, the idea that our universe is a simulation also entails determinate predictions: insofar as computerized simulations appear to require a *lattice-like* framework to account for distances between virtual objects and the progression of time, if our universe is a simulation, it too should be constructed on such a lattice framework. Finally, and most importantly, it has been shown that these lattices can emerge from quantum chromodynamics, and that the “edges” of these lattices should be observable to us in terms of an *energy limit* that can

⁶² See Y. Couder, A. Boudaoud, S. Protière, Julien Moukhtarb, E. Fort, “Walking droplets: a form of wave-particle duality at macroscopic level?”, *Europhysics News*, Vol. 41, No. 1 (2010), DOI: 10.1051/eprn/2010101.

⁶³ Again, see Hogan (2009) and Chown (2008).

be represented by energy particles.⁶⁴ Finally, Libertarian Compatibilism entails that we should observe specific *violations* of the quantum wave-function within human brains – violations that should still appear, in some way, to be the result of causally closed physical phenomena. Allow me to explain.

Consider my character in the Halo Simulation, “Master Chief.” If Halo Scientists investigated the behaviors of “Master Chief,” they would observe certain *propensities or tendencies* on “his” part. After all, I am controlling Master Chief, and my particular playing habits would display themselves within the simulation through Master Chief’s behaviors. Because observers in the Halo simulation would only have access to Master Chief’s propensities, they could not know how he would act before he actually acts. At most, they would be able to specify an empirically-verified probability distribution to model his future behavior. This probability distribution would, in laymen’s terms, specify something like this: “If Master Chief is in physical situation S , then there is a probability P that he will perform action A , a probability P^* that he will perform action A^* , etc.” Let us call the *probability function* that accurately describes my playing tendencies in Halo (tendencies that would reveal themselves to observers in the simulation) the “Master Chief Wave Function.”

Observers in the Halo simulation would say that the Master Chief Wave Function is an accurate description of Master Chief’s behavior in their physical reality, and that at every instant of observing his behavior, the Master Chief Wave Function will “collapse” into a determinate value (i.e., his actual behavior at that instant). The point of all this is that although things would appear this way to agents within the simulation, outside of

⁶⁴ Silas R. Beane, Zohreh Davoudi, Martin J. Savage, “Constraints on the Universe as a Numerical Simulation”, *arXiv: 1210.1847* [hep-ph] (2012), accessed on Nov 10, 2012 at <http://arxiv.org/abs/1210.1847>.

the simulation, I (a “nonphysical entity” from the perspective of the simulation) am choosing which value(s) the Master Chief Wave Function “collapses” to. I am doing this simply by moving a joystick around, causing the laser apparatus in my game machine to read one value (e.g., A^*) rather than another (e.g. A).

Interestingly, unlike the ordinary wave-function of quantum mechanics, the Master Chief Wave Function would not appear to “collapse” at random (according to the quantum wave-function). My individual propensities as Master Chief’s user (outside of the simulation) would manifest themselves *non-randomly* in his behavior. This means that, according to Halo Scientists, the physics governing Master Chief’s brain (supposing he has one in the simulation) would be slightly different than the physics of ordinary objects (like bullets). Whereas the quantum wave function would collapse randomly for all other physical objects, the physical mechanics of Master Chief’s brain (which result from my genuine free choices outside of the simulation) would be fundamentally *non-random*.

This leads us to the third unique, determinate empirical prediction mentioned earlier: Libertarian Compatibilism entails that if, and when, we are capable of studying quantum-mechanics in functioning human brains, the “usual” random collapse of the wave-function (which occurs for all other objects) should be subtly *violated* in non-random ways. Notice that it is perfectly consistent with Libertarian Compatibilism that these nonrandom effects could appear to have some “physical” explanation. For example, scientists in our world might discover that the human brain appears to have strange “feedback” mechanisms that appear to give rise to the slightly non-random elements of quantum collapse. Whatever “physical” explanations there might be for

these nonrandom elements, they would not undermine Libertarian Compatibilism, because Libertarian Compatibilism asserts that all of the “physical” events we experience are ultimately (partly) *enacted* by completely free choices in a higher-level frame of reference (outside of our physical reality). In other words, our free choices outside of physical reality could appear to have a “physical” explanation, just as Halo Scientists might create an elaborate “physical” explanation of the Master Chief Wave Function. The crucial thing to note is that Libertarian Compatibilism (A) *predicts* these nonrandom aspects of physics in the brain, and (B) explains whatever *physical appearances* those nonrandom aspects would have to observers inside the simulation (i.e. whatever special “quantum laws” scientists might construct to “physically” explain them). Whatever physical manifestations might appear to cause the nonrandom quantum effects would in fact be generated by the very same illusion that all the other phenomena were generated by. Thus, if we discover such nonrandom quantum effects in brains (however subtle), Libertarian Compatibilism will have made a true prediction that no other theory of quantum mechanics makes.

I contend, as such, that Libertarian Compatibilism makes three predictions that, jointly (or in combination), no other physical or metaphysical theory makes. Note that these three predictions may not be possible to verify or falsify for some time. Although the Holographic Principle is currently being tested, and it may be possible to test whether the observable universe is based on a lattice construct, Libertarian Compatibilism’s third prediction – that the quantum wave-function should be violated within human brains – may not be testable for some time. We are, after all, currently only able to test the predictions of quantum mechanics outside of the human brain,

using enormous particle colliders. Given that we cannot (and morally should not!) collide human brains together at high speeds to test how quantum-mechanics applies to them, how could we ever test Libertarian Compatibilism's empirical prediction? We would have to simulate human brains – perhaps in a vast neural network, in a supercomputer – and see whether those simulated-brains satisfy or do not satisfy traditional quantum-mechanics. Given that human, carbon-based neurons may somehow function differently than silicon-based processors (perhaps only organic, carbon-based neurons can “house” consciousness), we would first have to test the empirical adequacy of the created neural network (i.e. ensure that the neural-network simulation models every relevant aspect of organic, carbon-based neurons). Are such tests possible? If so, what would their results be? Only time will tell whether Libertarian Compatibilism's predictions will be verified or falsified. For now, at least, we have this: Libertarian Compatibilism accounts for numerous phenomena we presently experience – quantum indeterminacy, various philosophical problems, etc. – while making determinate physical predictions about what we should observe in future inquiry. Libertarian Compatibilism is therefore a powerful, live metaphysical theory with unique and definite empirical implications.

§4.8. Comparing Libertarian Compatibilism to Existing Interpretations of Quantum Mechanics

As we have seen, the Copenhagen, Everett/many-worlds, Bohmian, and Libertarian Compatibilist interpretations of quantum mechanics all make *unique* empirical predictions. Although none of their predictions are currently feasible to test, it is likely that such experiments will be feasible to perform in the future. Only time will tell, then,

which of the interpretations corresponds to reality. Until then, we must evaluate the various theories in terms of various desiderata commonly used to evaluate scientific and philosophical theories: explanatory power, simplicity, etc.

First, let us discuss the charge that Libertarian Compatibilism runs far more afoul of Occam's Razor than rival interpretations of quantum mechanics (Occam's Razor being the theoretical requirement not to posit entities and properties unnecessarily that do not (a) explain anything and/or (b) are not liable to empirical verification or falsification). I submit, to the contrary, that Libertarian Compatibilism does not posit entities unnecessarily, and that in fact the other interpretations of quantum mechanics *do*. To illustrate, consider the Copenhagen interpretation. Aside from arguably being incoherent in two respects, the Copenhagen Interpretation states that the process of taking "measurements" somehow lead us to observe the quantum wave-function collapse. How? Nobody knows. The Copenhagen interpretation thus posits some completely unexplained mechanism(s) through which "measurements" (a vaguely, possibly ill-defined concept within the theory) cause us to observe quantum collapse. This violates Occam's Razor. Now consider the Everett interpretation (i.e. the "many-worlds" interpretation). The Everett interpretation asserts the (completely unexplained) existence of a vast multiverse in which there are countless *conscious* variants of you, I, and everyone else. This has always been the main objection to the theory: while it seems to solve problems with the Copenhagen interpretation by simply defining "measurement" as "every possible physical state," it does so at an enormous ontological cost – namely, a vast array of universes equally "actual" as our own. This too violates Occam's Razor, at least if we can account for everything in our reality *without*

positing a vast array of equally “actualized” universes (which Libertarian Compatibilism accomplishes). Next, Bohmian Mechanics clearly violates the Razor as well: it posits hidden processes (or “variables”) that are unobservable in principle.

Now let us examine Libertarian Compatibilism. At first glance, it may seem to run even more afoul of Occam’s Razor than the previous theories, as it asserts (A) the existence of an (unexplained) physical multiverse (similar to the Everett interpretation), but also (B) *hidden processes* that somehow link all conscious choices and perceptions together, leading to a common experience of “quantum collapse”, and finally (C) *immaterial minds* outside of the physical order. In doing so, doesn’t Libertarian Compatibilism *combine* the ontological extravagances of each of the above theories into one, thereby making it a far worse violator of Occam’s Razor than any of the other theories? The surprising answer is no. For although Libertarian Compatibilism posits a great many entities, they are all *explained* and do important *explaining*.

Consider first the worry that Libertarian Compatibilism posits the same unexplained multiverse as the traditional Everett/many-worlds interpretation. Of the two theories, only the Everett interpretation leaves the multiverse unexplained. Libertarian Compatibilism actually explains the multiverse’s existence *in the simplest way possible* given our empirical evidence. Here is how. Suppose you wanted to build a bridge and you wanted to know whether a particular arch was necessary to make the bridge stand. How would you test this hypothesis? The answer is that you would build a *model* of the bridge and see whether the arch is necessary for making the bridge stand. Now turn to Libertarian Compatibilism. What does it *take* to model a reality like ours?

Libertarian Compatibilism gives us the answer, an empirically testable model of what it takes to create a reality like ours: MMORPGs. Insofar as MMORPGs reproduce all of the general features of our world – (A) a three-dimensional world of objects, (B) quantum indeterminacy and measurement problems, and (C) most of the philosophical problems we face – while structurally corresponding to Libertarian Compatibilism’s model of our reality, Libertarian Compatibilism seems to be the *only working model* we have ever created capable of *explaining* everything we experience. MMORPGs thus suggest that *in order to create a “world like ours”, there must be a “multiverse” of sorts – an informational array of possible pasts, presents, and futures.* Accordingly, Libertarian Compatibilism does not leave the multiverse unexplained. It *explains* the multiverse as a *necessary component* for creating any “world like ours” at all.

Indeed, it is worth dwelling on this a bit more. If you wanted to know how a bridge must be built in order to stand, you would create models of bridges and see which ones in fact stand. By analogy, if you wanted to see *what it takes* to create a world like ours (one which has the same “appearances”), you should try to create worlds like ours and see what components (i.e. materials, etc.) are necessary. This is exactly what we have in MMORPGs. In our own attempts, we see that in order to make a “world,” one must have two things: (A) hardware (i.e. a medium to read information in real-time), and (B) software, or “eternally” existing information comprising possible states of affairs. The only “world simulations” we have ever created all have dual processes: hardware and software. We simply don’t have *any* other working model. Libertarian Compatibilism does not leave the multiverse or mind-body dualism completely unexplained. Our experiences creating MMORPG simulations strongly, and

empirically, suggest that in order to make a “world like ours” – one in which observers experience “quantum indeterminacy”, the “explanatory gap” in the philosophy of mind, etc. – such a world *must be computationally erected upon a set of structures posited by Libertarian Compatibilism*: (A) “software”, i.e. an array of information comprising possible pasts, presents, and futures, and (B) “hardware”, i.e. a medium distinct from the software to run the software on (i.e. a computer console or consciousness).⁶⁵ As such, Libertarian Compatibilism provides the very first explanation of mind-body dualism. Insofar as existing MMORPGs simulate several key features of our reality, and the only way we know how to make an MMORPG is through a dualistic system containing (A) hardware (e.g. the game console), and (B) software (i.e. the game CD), our evidence strongly suggests that *one cannot create an anti-dualistic world*: all “worlds” need both “hardware” and “software.” Libertarian Compatibilism thus explains the very things – the multiverse, mind-body dualism, etc. – that rival theories

⁶⁵ This raises the obvious question: if every “world” must be erected upon some such *structure*, doesn’t it follow that the higher “level” of reality asserted by Libertarian Compatibilism (i.e. our consciousnesses) must be founded upon some such structure (viz., our consciousnesses must be explained by processes in a higher reference-frame – processes akin to MMORPGs’ “game consoles”)? But in that case wouldn’t it follow again that those processes (i.e. the processes in the higher-level frame of reference that give rise to *our* reality) must in turn be erected upon an even *higher* level structure; and on and on, *ad infinitum*? This is reminiscent, obviously, of the canonical worry regarding the famous myth that the world rests on the back of a turtle: namely, that if the world rests on a turtle, what does the turtle rest on? Although I cannot adequately address these questions here, let me say a few brief things. First, a leading theory in physical cosmology today – the theory of “eternal inflation” – states that our universe is embedded in an infinite hierarchy of universes which are eternally “birthing” new universes (for a thorough review of the theory of eternal inflation and evidence in favor of it, see Alan Guth, “Eternal Inflation and Its Implications”, *Journal of Physics A: Mathematical and Theoretical*, 40 (2007): 6811-6826). In other words, perhaps our best theory of physical cosmology today entails that “the world” in our frame of reference (i.e. the physical reality in our reference-frame) really is “turtles all the way down.” I see no reason not to think that all of reality has a similar structure. If every structure must be erected upon a structure, then reality may indeed have to be an infinite array of structures embedded in other structures, etc. This raises profoundly difficult questions on the notion of an “actual infinity” (see footnote 38). We must leave these issues for another day.

leave unexplained. It suggests that these sorts of objects are necessary for creating a “world like ours.”

Finally, Libertarian Compatibilism not only explains those things: it puts them to great use *in explaining*. First, whereas traditional interpretations of quantum mechanics simply attempt to model the “how” of quantum-mechanics (i.e. how quantum systems are indeterminate, measured, etc.), Libertarian Compatibilism explains *why* our world has the quantum features it does. Indeed, it gives us a model for how and why those features are produced: MMORPGs simulate the very “behind-the-scenes” processes necessary for observers in a “world” to experience quantum-effects. Indeed, I mean this point in the strongest way: to the extent that we understand the computational structures of the MMORPGs, we actually have a model of what is likely going on “behind the scenes” of our world (within whatever unobservable, higher-level reference-frame our universe is constructed). Given that we “see” the very same kinds of things as observers inside the simulated world of Halo, we have good abductive evidence that our entire world is constructed on a similar kind of structure (albeit a much more vast, far more elaborate one).

In addition to explaining all of these things, Libertarian Compatibilism explains the very existence of philosophical problems. It *predicts and explains* the kinds of questions philosophers ask about consciousness, time, personal identity, and free will. In *reproducing* the very philosophical problems we face in our reality, MMORPGs show that anyone in a Libertarian Compatibilist world would ask the *very philosophical questions we do*. Last, but certainly not least, Libertarian Compatibilism accomplishes that which has always been considered impossible: reconciling (A) genuine libertarian

free will, or free choice completely unconstrained by physical laws, with (B) the hypothesis that our physical world is completely “causally closed” (i.e. every physical event is either completely determined by previous events or the result of indeterministic laws, such as the equations of quantum mechanics). Libertarian Compatibilism shows how these things can be reconciled by way of MMORPGs. A world that appears completely causally closed from one frame-of-reference – a “physical” frame of reference *within* a world – can be generated by fully free choices in a *higher* reference-frame.

Now, of course, there is one momentous difference between the videogame analogy and Libertarian Compatibilism that I have set aside until now. The problem with the videogame analogy is that our choices outside of the simulated reality may themselves be completely determined by the physical laws of our world. In contrast, Libertarian Compatibilism hypothesizes that our consciousnesses are *entirely* free from determining physical (or perhaps nonphysical) laws, even in some higher frame-of-reference. Yet, what evidence do we have that this – perhaps the most crucial part of the theory – is true? Might a very close alternative theory be true instead: that every part of Libertarian Compatibilism is true, except that our nonphysical consciousnesses are in turn completely determined by “higher” nonphysical laws? My reply is simple (and, I hope, not too disappointing): I don’t think we can ever know either way. Just as the characters in my “Halo” videogame could never know whether the mechanisms outside of their simulated world (i.e. our choices, as their “users”) are completely free or determined by some further laws, the same is true with us. We can never know whether Libertarian Compatibilism is true – that is, whether our consciousnesses are

completely free to choose their way through the multiverse – or whether there are some higher nonphysical laws that constrain us. But even if we can't know this, we can at least say that we have some real evidence for the following disjunction: *either* (A) Libertarian Compatibilism is true, and our consciousness are entirely free to choose our paths through the multiverse, *or* (B) something very similar to Libertarian Compatibilism is true, but our consciousnesses are in turn fully determined by nonphysical laws (in a higher-level frame-of-reference) to which we can never obtain access. Both sides of this disjunction are equally fascinating. The possibility that we may never be able to obtain evidence favoring one hypothesis over the other may be disappointing – but it may well be the situation that we are simply consigned to (just as the characters in the “Halo world” can never know the truth about our physical world outside of their “reality”). Our best evidence may be that Libertarian Compatibilism may be true – no more, no less. While we undeniably would like to know for sure, perhaps knowing that we may have genuine free will – in the way that Libertarian Compatibilism posits – is the most we can ever hope for.

Before concluding, there is one final issue worth addressing. Our conscious experience of time's passing feels “smooth” – an unbroken “flow” into the future. Libertarian Compatibilism, on the other hand, holds that our experience of the passage of time consists of our consciousnesses “hopping” across different branches of a vast multiverse. So when do these “hoppings” occur – at every instant, or at specific “moments” determined by lawfully regulated events in a nonphysical realm (the laws of which we do not know)? Although we may be necessarily ignorant of whatever mechanisms or laws govern the nonphysical realm of consciousness that Libertarian

Compatibilism posits, I believe that the above questions have a determinate answer from our frame of reference. Once again, I will explain by analogy to online videogames. The simulated “reality” of a videogame is simply a string of information that the game console is reading off of the game CD. The strings of information on the game CD define discrete “decision-points” – points at which the player of the game can in fact make a choice. Since the string of information the game machine plays at any given instant is comprised by that discrete series of information, agents within the game environment – if they were to inspect their world’s physics from their frame of reference – would, in principle, perceive discrete packets of information (i.e. quanta) representing discrete instants at which events occur and “choices” are made. Again, from their perspective, all of these events may well appear to be “fully determined” – but in reality, the outside user would have the capacity to make a choice at any of those discrete instants of information. In short, Libertarian Compatibilism suggests that we should indeed understand quantum events in our world as discrete, instantaneous events at which our nonphysical consciousnesses choose which branch of many possible subsequent instants are experienced next.⁶⁶ Finally, I believe that the “playing” of these instants, one after another, gives rise to the *illusion* of smooth forward flow over time – just as in a videogame. For example, when I play the game “Halo,” its events ordinarily seem smooth in the same way that conscious experience seems smooth – when all they really are is a game-player scanning discrete packets of information. Insofar as discrete quanta of information can create this illusion of “smoothness” in a game environment,

⁶⁶ The idea that “hoppings” from branch to branch of the multiverse occur at these instants coheres well with the phenomenon of *decoherence* in quantum physics, where each “splitting” of the multiverse into different branches is given a quantum-mechanical description. See Zurek (2003), Schlosshauer (2004), and especially Bousso and Susskind (2011).

Forthcoming in *The Philosophical Forum*

Libertarian Compatibilism suggests that our experience of the smooth flow of consciousness is a similar kind of illusion generated by our consciousnesses, as nonphysical media, reading the discrete quanta that comprise our physical reality.

Conclusion

The theory of free will explored here – Libertarian Compatibilism – is fantastic and may turn out to be false. The important thing, however, is that it may well be *true*. Several serious philosophical and scientific hypotheses – all of which have some real evidence in their favor – jointly suggest the theory is, at the present time, a serious epistemic possibility. Further, Libertarian Compatibilism is a powerful theory: one that explains quantum mechanics as well as the very appearance of an array of notorious philosophical problems in the philosophy of mind, time, personal identity, and free will. Finally, Libertarian Compatibilism makes three determinate – and unique – empirical predictions: predictions that may, in the future, be either verified or falsified by the physical sciences. At a time when the philosophical debate about free will has arguably stagnated, with the familiar positions – compatibilism, incompatibilism, and libertarianism – repeatedly trotted out, given new defenses of, new objections raised to, and the process then repeating all over again, it is my hope that the new theory of free will defended will be seen, at the very least, as what it is: a provocative, fascinating, *possibly true*, breath of fresh air.