

The Twin Vantage Point Paradox: A Thought Experiment

Baland Jalal

University of Cambridge

Vilayanur S. Ramachandran

University of California at San Diego

As noted by Erwin Schrödinger, in the world of physics the subjective vantage point does not exist. Indeed, physics requires that the conscious “self” who experiences the world be banished from reality. This article highlights this schism running right through the heart of reality using a thought experiment. The reader imagines himself as a twin being asked by a “super-scientist”: “If I travel back in time and swap you at birth with your twin (now an adult who lives in an adjacent city), and then return to the present, who would you prefer that I torture here and now — the person standing in front of me — or the person in the other city?” To this question (ethical considerations aside), you might answer, “torture the twin in front of you” — that is, you would assume given the swap that you would be existentially continuous with the twin now living in the adjacent city and vice versa. This culminates in a paradox: while for you, such a swap matters crucially, according to science, the final state of the universe is the same regardless of the swap. This paradox does not exist from a scientific objective view of reality — only from a subjective existential point of view. To our knowledge, this issue of continuity of form and function, and the fact that the physical universe is the same irrespective of the swap, is not part of many published thought experiments exploring personal identity.

Keywords: subjective vantage point, physicist’s worldview, twin replacement

“We are such stuff as dreams are made on, and our little life is rounded with a sleep.”
 William Shakespeare, 1611, *The Tempest*

In physics, there is an elimination of the subjective vantage point — or the “self” who experiences the world. The conscious “self” is merely neural activity and nothing more; colors do not exist, only wave lengths; there is no such thing as warmth

We are grateful to Raymond Russ, editor-in-chief, for his assistance with editing this manuscript. Correspondence concerning this article should be addressed to Baland Jalal, Department of Psychiatry, Trinity College, University of Cambridge, CB2 1TQ, United Kingdom. Email: bj272@cam.ac.uk

or cold, only kinetic activity of molecules; no pitch, only frequency. Yet for me, my subjective “I” is my only access to the world, the center of all things; as raised by the Upanishads, an ancient Indian philosophical text dating back to second millennium BC: “If we halt in subjective mind, we see the objective world as a mere dream or a vision of our conscious subjective activity” (Aurobindo, 2004, p. 405). Many years later, Schrödinger (1944/1992) wrote about the “principle of objectivation”:

By this [i.e., the principle of objectivation] I mean what is also frequently called the “hypothesis of the real world” around us. . . . Without being aware of it and without being rigorously systematic about it, we exclude the Subject of Cognizance from the domain of nature we endeavor to understand. We step with our own person back into the part of an onlooker who does not belong to the world, which by this very procedure becomes an objective world. (p. 118)

Elsewhere he wrote:

So we are faced with the following remarkable situation. While the stuff from which our world picture is built is yielded exclusively from the sense organs as organs of the mind, so that every man’s world picture is and always remains a construct of his mind and cannot be proved to have any other existence, yet the conscious mind itself remains a stranger within that construct, it has no living space in it, you can spot it nowhere in space. (p. 122)

According to these passages, only a detached external observer exists in physics — the subjective “self” simply does not. Your “I” or mine is merely a brain out of billions of brains; a byproduct of evolution, with no privileged access to the world. On the other hand, for me it appears as if a tiny space of the Minkowski space–time diagram¹ is illuminated by my conscious experience. We shall argue, following Schrödinger’s lead, that this is a rift running through the heart of reality, not merely a linguistic issue.

The Twin Vantage Point Paradox

A pair of monozygotic twins is born in San Diego. One of these twins is you. (For the sake of argument, we will make the simplifying assumption that you and your twin are identical down to every detail [but as we shall see later, this assumption is not critical for the argument].) You and your twin are separated at birth; you remain in San Diego while your twin is brought up in Los Angeles (LA). When you are of age 35, a time-traveling “super-scientist” approaches you, and asks, “if I travel back in time and swap you at birth (so who is now “you” would be the one living in LA and your twin is in San Diego), and then come back to the present, who would you prefer that I torture right this instant, the person in LA or the one in San Diego, if

¹Hermann Minkowski (1909) famously introduced the Minkowski diagram that illustrates the properties of the four-dimensional space–time manifold of Einstein’s special relativity theory.

you had to choose one?” To this question, given the natural inclination to avoid pain (ethical considerations aside), you might answer, “torture the person here and now in San Diego, because that other person in LA is existentially continuous with me, in fact he is me” (in the sense that the child who originally grew to be the adult you are now talking to, would have as a result of the swap become the person in LA); on the other hand, if the super-scientist asks you the same question but with no swap involved, you would answer, “torture my twin in LA.”

As we shall see, it is difficult to defend this position from a strictly scientific point of view, since the final state of the universe is the same regardless of the swap. This is so because you and your twin's bodies — atoms and molecules — are replaced in cycles (i.e., undergo complete metabolic turnover). With each cycle, the molecules of your bodies and brains are replaced with a new set of molecules — you literally incarnate a new physical body. While you encounter food and air molecules in San Diego forming the basis of your bodily cells (let us call these “SD atoms”), your twin encounters food and air molecules in LA (let us call these “LA atoms”). So your body is now made up of SD atoms and your twin's, LA atoms. This interaction of your physical bodies with the environment over the span of 35 years, results in two bodies, two separate twins. Your present bodies (at the age of 35) are merely replicas, or copies of your previous bodies (last time all your cells were replaced). From this follows that the fetuses of origin, or first bodies, serve as mere templates for the entire sequence of subsequent bodies (i.e., each full-body and brain molecule replacement). Now imagine that you and your twin had been exchanged at birth. Then you would have been exposed to LA atoms and your twin to SD atoms. But regardless of this exchange, the final state of the universe would be exactly the same, had the swap not taken place. Put differently, given the identical nature of you and your twin's genetic make-up and the ongoing replacement/renewal of your physical bodies (i.e., the hardware), according to science, the ultimate state of the universe is unaltered, regardless of the swap (see also Ramachandran, 1980). [Even though, the final state was arrived at via different trajectories in space-time, the end result is the same — not merely identical with what would have occurred had the swap not taken place.] What also follows from this, is that since the physical world is unaffected regardless of the swap, your personal existence would also remain unaffected. Yet, for you, the swap matters crucially! Thus, what the physicist's worldview — his reflex response to this riddle — does not take into account, is the continuity of conscious experience, the subjective “self,” or what in metaphysics would be regarded as the soul.

Another way to state this “twin vantage point paradox” would be to imagine that your twin was brought up in LA under less fortunate circumstances. You, on the other hand, have a life of wealth and privilege in San Diego. Since your genetic make-up is identical, the environment alone made this difference. Now suppose a super-scientist was to travel back in time and swap you at birth. Given the above argument (i.e., at least one metabolic replacement of your atoms has occurred since

the separation), you would now exist as your twin, and experience all those unfortunate circumstances that once belonged exclusively to him (Ramachandran, 1980). [Although you would most likely protest to such a swap, you would be unable to logically defend the swap; since the world remains unchanged physically.]

Another factor that is relevant to this “twin paradox” is the question of whether the replacement of the atoms is done suddenly rather than gradually. Suppose a super-scientist were to create a being identical to you, down to every detail. The super-scientist then places your manufactured “twin” (or replica) in a room, similar to the one you are sitting in. If he then asks you, “Who would you prefer that I torture here and now, you or your ‘twin’ in the other room?” you might answer (again, ethical considerations aside), “torture the guy in the other room, because I won’t actually feel his pain.” If the super-scientist then were to reveal that he would kill you and asked, “Would you mind that I torture your twin after your death?” you may answer, “although I wish him the best, since I won’t actually feel his pain, I don’t particularly mind what you will do to him.” Yet paradoxically, from a scientific view of the world, if the super-scientist kills you and then sits your “twin” in your chair, this would be logically equivalent to replacing all the atoms in your body; indeed, it would be no different from asking if you yourself would mind being tortured a year from now — despite the fact that you would have been replaced several times by then. (The fact that the atomic replacement is done suddenly rather than gradually, is irrelevant as the final result of the replacement is the same.) Again, the scientific worldview does not take into account the continuity of conscious experience — existential continuity is unaffected by the atomic replacement. It also follows that if the super-scientist now tortures your twin, “you” would (according to this argument) feel his pain, which means you ought to be as concerned about your twin’s future well-being as your own. While you will not feel his pain as long as you are alive, the instant he sits your “twin” in your chair after first killing you, arguably, you will feel his pain (Ramachandran, 1980). There is simply no empirical test that can resolve this paradox, which in some eyes would render it meaningless. But as conscious human beings we cannot help but wonder.

In the above example, if you think about it, there are three ways in which you (A) and your twin [replica] (B) sitting in the adjacent room, could be swapped by the super-scientist: (1) he could make you exchange atoms one by one (like the ship of Theseus) across the room, and you would not notice an interruption of conscious experience [even if there was such an interruption]. (2) He could pulverize both of you and move your (A’s) atoms to B’s room and vice versa, and then resurrect you. (3) You simply walk across the room passing each other, swapping place; in this third case, most people would obviously pick the guy who was earlier sitting in the chair you now occupy, to be tortured (if you had to choose).

Dilemmas such as these suggest, perhaps, that intuitions about numerosity and identity etc., are very primitive and do not apply in the same manner to objects,

functions, and especially minds. This could lead to the cynical inference that these questions are meaningless and emerge from the limited resolving power of ordinary language (the Wittgenstein school), as opposed to genuine metaphysical questions outside the scope of science.

Notice that in the first and second example above, the atoms of your brain have been changed completely. Therefore, the final state of the universe after the swap is different; unless you evoke the quasi metaphysical notion of “identity of indistinguishable particles.” Contrast this with the “baby swap/time travel” thought experiment above, in which several cycles of replacement have occurred, with the net result that the final state of the universe is unchanged — what is changed is your world-line in the Minkowski space-time diagram; in other words, in this case, the only difference between you and your twin, is the historical trajectory/functional continuity with the past, but not the end state of the universe.

As noted by Parfit (1984), similar paradoxes arise in split brain patients, in whom the corpus callosum connecting the two hemispheres of the human brain has been cut. We now introduce a novel paradox. We know that each hemisphere of the brain is fully autonomous, and this can be tested empirically in the laboratory (in split brain patients) to show that both hemispheres are independently “conscious.” Each hemisphere has laid down the memories of a lifetime, and enjoys “personhood” to the same extent as a “normal” conscious human being. Yet puzzlingly — this raises a paradox overlooked in most accounts — each hemisphere can claim continuity with the single agent existing in the past, asserting a unique and private relationship with that person.

Now imagine if I were to present a question (visually) in the right visual field of a “normal” healthy person; the left hemisphere would understand the question and reply. Imagine that person is you. Let us say I ask, “Would you mind if I poke either one of your hands with a hot rod (that will leave no tissue damage), and in exchange give you 100 dollars?” You might answer, “of course I would mind.” But if I (the experimenter) ask you, “I am now going to anesthetize your corpus callosum, and then burn your left hand, would you mind?” you may answer, “no, in that case I won’t mind.” Since you know your anatomy, you realize that the “you” that is speaking to the experimenter would no longer experience the pain administered to the left hand, because the left hemisphere, in the absence of a functioning corpus callosum, does not have access to pain delivered in the left hand.² Yet this poses a conundrum, because just moments ago, “you” (in the left hemisphere) were the same person as the “other” one (in the right hemisphere). But because of the “amoeba-like” split, you are no longer worried about the other hemisphere’s pain; that is, until the anesthetic wears off and you are one person again.

² Technically there are certain ipsilateral pain pathways as well, but this does not affect the logic of our argument.

Lastly, we have the even more puzzling conundrum of what would happen if atoms in your brain (A) were replaced, not by a single new atom for each atom removed but by two? You would then have been gradually replaced with two identical beings, side-by-side (B and C). This raises the question — do you continue in both beings or just in one? It cannot be both because then if you put one being in a super-hot room and the other being in an ice-cold room — “you” would experience both hot and cold simultaneously — which makes no logical sense. But if you continue in only one (B), who or what decides whether you are B or C?

Concluding Remarks

We believe our “twin vantage point paradox” raises the most fundamental question you can ask as a sentient human being; namely, the “here and now” of conscious awareness. Why and how is your mind “loyal” to the particular brain currently associated with it? It should be emphasized that there is no paradox from a strictly scientific objective third-person view of reality. For you (the “reader” who imagines himself as the twin) being approached by the super-scientist, such a swap is critical — and determines whether you want yourself or your twin to be tortured.³ However, from an objective third-person view of reality, such swapping of the twins in the cradle makes no difference at all, since the final state of the universe is unchanged regardless of the swap. The physicist’s worldview does not take into account the relationship between the “objective” functional continuity of informational states encoded by neurons, with the sense of continuity of conscious subjective experience. And since the subjective “self” does not exist in the physicist’s worldview, the swap makes no difference. Thus it is only a paradox if you imagine yourself as one of the twins. It arises only in metaphysics, and not in science.

For the sake of simplicity, we assumed that the twins were identical down to every detail. However, mental states fluctuate from moment to moment (as mentioned, so do physical bodies, including neurons). Such fluctuation, however, does not affect the continuity of conscious experience, or lead to a “jerky” existence, and is therefore not relevant to the argument being put forth.

We emphasize that this doctrine, although it has a “soul-like” ring to it, is different from the orthodox religious conception of a soul as an entity divorced from matter. Also, our paradox reminds us of the ship of Theseus (continuous

³The ethical question of whether you would actually “prefer” yourself or your twin being tortured — even if you are a “Gandhi-like” person — is irrelevant to the argument being presented. The fact that such a swap is crucial to your subjective experience (torture or no torture), is what is critical, yet not accounted for by a scientific worldview, which only takes into account the third-person view and the final state of the universe.

replacement of brain atoms = replacement of ship boards).⁴ The structural continuity aspect of reality and associated functional continuity (the crux of the Theseus paradox), plus the fact that in our thought experiment the final state of the universe remains unchanged in spite of the swap, are not part of many thought experiments exploring personal identity (e.g., Lyon, 1980; Parfit, 1971, 1982, 1984; Shoemaker, 1963, 1970, 1984; Vesey, 1973; Williams, 1970, 1973). Most such thought experiments assume tacitly that only the information carried by spatiotemporal patterns of neural activity and not the physical stuff (or its continuity) is relevant; thereby begging the very (metaphysical) question that it is meant to address. Central to our argument is the distinction we make between empirical identity — in the ship of Theseus sense — and existential identity; the brief “illumination” that the flame of your consciousness brings to a bleak, sterile and inhospitable world.

Thought experiments exploring personal identity are not new in philosophy (e.g., Campbell, 2004; Locke, 1694/1975; Olson, 1999; Parfit, 1971, 1982, 1984; Schechtman, 2014; Shoemaker, 1963, 1970, 1984; Vesey, 1973; Williams, 1970, 1973). In particular, Parfit’s famous classic, *Reasons and Persons*,⁵ considers similar paradoxes that arise in hypothetical brain transplants. However, there appears to be a reluctance in the field to make the distinction between the purely semantic question, of whether you would call the entity participating in the swap, the same or not (should be a matter of definition), versus the genuine ontological question which, we argue, arises in our thought experiment. Indeed, the vast majority of scholars begin with the physical world as an axiom beyond question, and the paradoxes like the one we raise, would automatically be discarded as purely linguistic riddles.

The Simultaneity of Conscious Experience

Our essay considers only paradoxes of the continuity of consciousness following destruction and resurrection of brains; we barely touch on the question of simultaneity. In particular, if continuity of function is all that matters, then what happens if a scientist replaces your brain (gradually, if necessary — as in the ship of Theseus) with two atoms for each of yours, so you are cloned into two co-existing twins. Would you continue in both twins in parallel and — if not — what determines

⁴If we begin with the ship of Theseus floating on the ocean, and replace all the planks one by one, so the ship never sinks, until you have replaced all the planks: Is this still the ship of Theseus or not? Or consider another analogy: if a candle (A) is used to light another candle (B), the new candle is made up of completely new material. But is the flame a new flame or the same as the old flame? And what about the illumination provided by the new flame? The insolvable nature of such questions makes them irresistible to the philosophers who wrestle with them.

which brain you continue in? The two minds can't be one — as asserted by mystics because they would raise the question of what would happen, if you put twin B on hot coals and twin C on a glacier? Is there a global consciousness that is simultaneously aware of heat and cold!? A partial impediment to our thinking about these issues, arises from our tendency to unconsciously think of minds as being a sort of “halo” that's spatially localized in a brain. We prefer the analogy of clocks measuring time. The passage of time doesn't depend on (nor is localized in) a particular clock; but the clock registers time nonetheless. Time is equivalent to consciousness and brain to clock.

References

- Aurobindo, S. (2004). *The Upanishads — I: Isha upanishad*. Pondicherry: Sri Aurobindo Publication Department.
- Campbell, S. (2004). Rapid psychological change. *Analysis*, 64, 256–264.
- Locke, J. (1975). *An essay concerning human understanding* (P. Nidditch, Ed.). Oxford: Oxford University Press. (Originally published 1694)
- Lyon, A. (1980). On remaining the same person. *Philosophy*, 55, 167–182.
- Minkowski, H. (1909). Raum und zeit. *Physikalische Zeitschrift*, 10, 104–111.
- Olson, E. T. (1999). *The human animal: Personal identity without psychology*. New York: Oxford University Press.
- Parfit, D. (1971). Personal identity. *The Philosophical Review*, 80, 3–27.
- Parfit, D. (1982). Personal identity and rationality. *Synthese*, 53, 227–241.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.
- Ramachandran, V. S. (1980). Twins, split brains and personal identity. In B.D. Josephson and V.S. Ramachandran (Eds.), *Consciousness and the physical world* (pp. 139–163). New York: Pergamon Press.
- Schechtman, M. (2014). *Staying alive*. Oxford: Oxford University Press.
- Schrödinger, E. (1992). *What is life? With mind and matter and autobiographical sketches*. Cambridge: Cambridge University Press. (Originally published 1944)
- Shoemaker, S. (1963). *Self-knowledge and self-identity*. Ithaca: Cornell University Press.
- Shoemaker, S. (1970). Persons and their pasts. *American Philosophical Quarterly*, 7, 269–285.
- Shoemaker, S. (1984). A materialist's account. In S. Shoemaker and R. Swinburne (Eds.), *Personal identity* (pp. 67–132). Oxford: Blackwell.
- Vesey, G. (1973). *Personal identity*. Milton Keynes: Open University Press.
- Williams, B. (1970). The self and the future. *The Philosophical Review*, 79, 161–180.
- Williams, B. (1973). *Problems of the self*. Cambridge: Cambridge University Press.