

# The Stream of Consciousness and Personal Identity

Keith Wiley

Board, Carboncopies.org; Fellow, The Brain Preservation Foundation  
[kwiley@keithwiley.com](mailto:kwiley@keithwiley.com), <http://keithwiley.com>

Aug. 10, 2019

## Abstract

This article marks the five-year anniversary of the publication of my book, *A Taxonomy and Metaphysics of Mind-Uploading*, and is loosely a celebration of the past five years. I first discuss the challenges of discovering and debating metaphysical truths. I then tackle an argument against discontinuous mind uploading, such as via scan-and-copy, based on the claim that it fails to preserve a purported temporal stream of consciousness, thereby losing the original metaphysical personal identity and invoking a replacement identity in the upload. This article investigates the stream of consciousness claim from several angles and concludes that it is not only unsubstantiated in theory, but also incompatible with real-world accounts of how people treat one another in existing medical cases that should, by all accounts, destroy the stream of consciousness. As an alternative, this article suggests that personal identity is better characterized by a person's memories in the synchronic specious present, with no requirement of diachronic psychological or phenomenal properties.

## Introduction

This article loosely coincides with the five-year anniversary of the publication of my book, *A Taxonomy and Metaphysics of Mind-Uploading*, in which I present a wide range of thought experiments about mind uploading and then convey my preferred model of identity, commonly known as branching identity. After publication of the book, I published a paper in the *Journal of Consciousness Studies* [Wiley & Koene 2016], coauthored with my long-time colleague, the neuroscientist Randal Koene. Dr. Koene is the founder of Carboncopies.org, which promotes philosophy and neuroscience of mind uploading. Randal further introduced me to Kenneth Hayworth, co-founder and president of The Brain Preservation Foundation, where I have written about next generation techniques for preserving brains that go far beyond the cryogenic and plastinated methods developed in the twentieth century [McIntyre and Fahy 2015]. I have greatly enjoyed presenting my ideas about the nature of metaphysical personal identity, with emphasis on its application to mind uploading and the underappreciated theory of branching identity [Cerullo 2015, Wiley 2014]. Regardless of if or when mind uploading ever becomes practical, it nevertheless provides an excellent sandbox for testing concepts of consciousness and identity.

Personal identity is a challenging topic to tackle with any intent of objective truth discovery. Metaphysical truths (if there are such things) are more difficult to pin down than physical truths because much of metaphysics is not amenable to empirical physical experimentation. Karl Popper initially gave us our nearly universal reliance on falsifiability as the demarcation criterion separating science from metaphysics—and from supernatural or religious beliefs for that matter [Popper 1962]. If a theory offers the notion of its own falsification by empirical experiment, even if that falsification has not yet been produced, then the theory is scientific, so explained Popper. Otherwise, the theory resides in philosophy,

metaphysics or faith. Thomas Kuhn, however, questioned whether empirical methods can offer us objective tools of reality discovery, since we can't help but interpret observations from the vantage point of our preexisting models of reality. If two people hold different models to begin with, then even unvarnished reality laid bare before their very eyes may not convince them who is right. Kuhn called this problem *incommensurability* [Kuhn 1962]. Not to jump too far ahead, but people often fail to convince one another of their preferred models of personal identity with respect to mind uploading because all conclusions are interpreted as supporting each person's preexisting model.

Assuming Kuhn left us any hope of forward progress, then once in the realm of metaphysics, Popper's falsification demarcation leaves us initially rudderless in weighing philosophical and metaphysical theories; it merely demarcates the scientific questions from the nonscientific ones. This is the question of metaphysical epistemology: how should we separate good metaphysical theories from bad ones to determine metaphysical truth? Popper believed that philosophy proceeds by the production of conjectures which are subjected to as many refutations as may be conceived. Conjectures that withstand this gauntlet become canon. We see this practice regularly in philosophical debates in the form of thought experiments, in which philosophical proposals are subjected to whatever criticism the proponent or broader community can devise. I personally proceed by seeking and driving out paradoxes and circular arguments in my metaphysical ponderings. Given my strong leaning toward honoring first person perspectives regarding consciousness, I see immediate paradoxes in proposals where people or uploads are regarded in ways that contradict their first-person points of view, or in scenarios in which multiple people with essentially identical points of view are nevertheless externally assigned differing identity labels. We will see another example in this article in which the neurological underpinnings of any identity threshold to a stream of consciousness are incompatible with real world medical cases. Discovered paradoxes are, to me, one of the strongest warnings of a poor metaphysical theory, and this approach has led me to some stark disagreements with others concerning the nature of metaphysical identity, primarily in my advocacy for branching identity. Consequently, explaining this theory of identity has become the bulk of my contribution to mind uploading philosophy, if there is such a thing.

Given metaphysics' lack of empirical experiments, as the physical sciences offer us, we are often led by our *intuitions* to locations on the philosophical landscape which we subsequently perceive as rather obvious conclusions, even as others are drawn by their own intuitions to different hills on the same landscape. Philosophers, both professional and otherwise, often come across in their writing as positively confounded—practically incredulous—that anyone could disagree with their concluded model of reality. I'm likely similarly guilty, as this article may bear out. It is perplexing that the following three statements should seemingly all be true:

- Metaphysics is unsolvable with the same rigor as the physical sciences, as Popper specifically demarcated.
- Most of us hold strong convictions concerning our positions on metaphysical questions.
- Our convictions are frequently incompatible with the convictions of others.

From these three observations, we may conclude that either some people are right and others are profoundly wrong on matters of metaphysics, or alternatively, we are all wrong about the very nature of metaphysics to feel that it contains genuine truths—or falsehoods—in the first place. Others have noted this problem about metaphysics before, namely that we each hold prior assumptions about reality that influence where our intuition ultimately leads us [Dainton & Bayne 2005]. Richard H. Wiley emphasizes the role of noise at every step along the way in matters of both personal truth discovery and conveyance of one's understanding of truth to others [Wiley, R. 2015]. Such pervasive noise not only confounds our search, but renders it essentially impossible for two people to ever truly perfectly understand one another. One possible explanation for how the three incompatible circumstances above come into being is

indicated by Kuhn's incommensurability: we are all constantly judging one another's hypotheses from different vantage points that are not as objective as we would hope. Another explanation is that our conviction about our individual philosophies originates not from methodical and logical induction, which both Hume and Popper rejected anyway, but from evolutionary forces, which produce methods of mental world modeling under no requirement to align with external truths, only with survival and reproduction.

Despite the recognized challenges in discovering—or arguably in choosing—metaphysical truths, we engage in it anyway. We all build mental models of the world, including those aspects of the world that are not physical and not empirically approachable, such as our own and others' minds (aka *theory of mind*). Starting with my book, I have spent the last five years attempting to sharpen the exposition of my particular bag of ideas on the nature of the mind and personal identity, namely concerning branching identity. This article is, to some extent, a celebration of the effects of the book's acceptance on my life over the last five years, of the intellectual colleagues I have met, and of any small public recognition that I or my ideas may have gained.

## The stream of consciousness

In this article, I discuss a common argument against spatially and/or temporally discontinuous procedures for mind uploading based on their apparent failure to preserve a proposed metaphysical phenomenon frequently described as a *stream of consciousness* (SoC). If a mind uploading procedure is deemed to lose track of this diachronic (aka across time) stream, then this theory judges the outcome as a death of the original identity with some new identity invoked as a replacement. This new identity is often referred to as a *copy* in popular vernacular, and in a disparaging way: to be a copy is judged an undesirable fate and indicates a failure to survive the procedure.

At first glance, we might assume that the SoC refers to the stream of waking thoughts of which we are specifically aware, if for no better reason than that that is precisely how the term was originally introduced by William James [James 1890]. It is his term after all, so he should get first call as to its meaning. But forbidding a breakage in that sort of stream, out of concern for loss of identity, is immediately confounded by examples such as sleep, fainting, general anesthesia, etc., sometimes known as the *bridge problem*. Clearly, that stream breaks on a regular basis without concern that identity is constantly lost and invoked anew, or at least without concern that such a model of identity is in any way disturbing.

Despite the term's unambiguous original intent, *stream of consciousness proponents* (SCPs) generally respond to counterexamples such as sleep by either seeking a bridging criterion to handle gaps or by redefining the SoC to represent a different concept deemed likely to persist across gaps in waking consciousness: any and all psychological traits, no matter how deeply subconscious or indirect in their causal influence upon our aware experiences. Furthermore, such an abstract stream is widely admitted to arise from some basal neural activity such that we may focus our examination not on intangible mental streams but on physical neural activity. The recast claim is that, should we break this foundational stream of neural activity, obviously breaking its associated mental features as a result, then the purported SoC will also break (which is not a particularly contentious observation), and that the apparent consequence is that metaphysical identity vanishes for good (this is the contentious claim). Similarly, should the neural activity restart at a later time, with all its previous mental features reestablished, the argument claims that the original identity does not reappear or continue, as would otherwise occur after presumably weaker gaps, such as sleep. Rather, the original identity remains permanently lost, while a new identity is invoked in its place. This is the bridge problem again, which I discuss below. Since this outcome is often characterized as a form of death with replacement by a different person, such mind uploading procedures

are generally judged as a lack of survival and consequently as an undesirable outcome. Dainton and Bayne solve the bridge problem by focusing on the potential for a persistent stream of conscious experience, even if unrealized in actuality [Dainton & Bayne 2005]. On the other hand, I conceive of identity in a radically different way than Dainton and Bayne, focusing not on a diachronic experiential stream, potential or otherwise, but rather on a speciously present (aka synchronic) memory of oneself as a person with a cohesive identity.

Although SCPs frequently argue against discontinuous mind uploading, they often accept the notion of mind uploading in principle, but insist upon a procedure that on some analysis is entrusted to preserve this stream. For example, consider the two classic mind uploading procedures known as *scan-and-copy* and *gradual in-place replacement*. In scan-and-copy, a statically preserved brain (cryogenic or plastinated) is sectioned into two dimensional slices, with each slice independently imaged, the overall three dimensional structure (aka the *connectome*) inferred from those slice images, the brain recreated either in simulation or as a physical artifact, and finally restarted. We are making impressive inroads on every aspect of this procedure and it may evolve from hypothetical thought experiment to practical viability within some foreseeable future. As described, scan-and-copy is discontinuous in four ways:

- Spatially: the upload is constructed in a new location relative to the brain.
- Materially: the upload is fully constructed from new matter before neural processing resumes.
- Temporally: neural activity enters stasis for some time period before later restarting in the upload.
- Functionally: neural activity comes to a global cessation and then restarts from scratch.

SCPs generally reject scan-and-copy for apparently breaking the diachronic conscious stream due to one or more of the four discontinuities listed above, thereby destroying the original identity and somehow invoking a new identity in the upload. On the other hand, gradual in-place replacement mind uploading, in which each neuron is steadily replaced within the head by a microscopic artificial neural prosthetic (a procedure that is essentially impossible by even the dreamiest technology at the current time), is often proposed as a method that successfully preserves one's SoC and associated personal identity. Gradual in-place replacement exhibits none of the four discontinuities described above. Although the upload's substrate is not the same as the brain, and therefore spatial translation and material replacement occur, these changes are not discontinuous. Likewise, although at an earlier time processing occurs in the brain and at a later time it occurs in the upload, the change is not discontinuous but rather piecemeal, for whatever benefit that may ostensibly offer.

Other hypothetical procedures are discontinuous on some subset of the four discontinuities listed. For example, cryonics is temporally and functionally discontinuous but spatially and materially continuous, since it involves reviving the original brain, while instantaneous teleportation involving only transmitted pattern information, but utilizing new matter, is the opposite of cryonics: temporally and functionally continuous but spatially and materially discontinuous. Teleportation involving a matter beam of some sort is only spatially discontinuous. The whole point of my *JoCS* paper with Koene, unambiguously titled *The Fallacy of Favoring Gradual Replacement Mind Uploading Over Scan-and-Copy*, is to thoroughly dispel the popular position that gradual replacement should be seen metaphysically favorably from the ongoing debate about mind uploading, i.e., to show that it is an irrational preference upon careful analysis. I urge critical readers to read it in full of course.

This article considers the SoC claim against discontinuous mind uploading from two approaches. First, I discuss the related proposal that spatial continuity is of importance on the claim that successful preservation of the purported SoC should hinge on body (or at least brain) spatial continuity. I then explore the popular feature of the SoC characterized as its apparent reliance on a never-ending dynamic metaphysical process (cognition) that arises from an underlying never-ending dynamic neurological process (neural firing activity).

## Spatial continuity

One argument that appeals to a certain intuition for why personal identity fails to preserve in a discontinuous mind uploading scenario involves a violation of spatial continuity. How does identity undergo a discontinuous transfer through space from inside a person's head to the upload's new brain or computer [Corabi & Schneider 2012]? It is as though some people visualize their identity as a literal homunculus, physically residing inside their head, that must travel across the room the way a physical bird flies through the air, ultimately landing in whatever physical system will instantiate the uploaded brain scan data so as to preserve (aka transfer) the person's identity from one location to another. Since it is unclear how such literal spatial travel of this identity object could be accomplished in a discontinuous procedure, the procedure is rejected as a failure to preserve identity. I don't intend this description to sound sarcastic. It is a genuine attempt on my part to comprehend the proposed claim. Contrarily, I don't believe metaphysical identities have physical properties in the first place, including the property of spatial location, so the entire concept of transferring identity from a brain to an upload's new substrate feels like a category error to me. My book goes into great detail on this issue.

Below, I address this concern of spatial transfer via an argument I have not considered before, based on discrepancies between our visual experience of our location and our true location. But first, both my book and my *JoCS* paper with Koene have already challenged the spatial-continuity argument based on the flawed expectation that gradual in-place replacement mind uploading (which, recall, SCPs often accept to preserve identity) would not itself involve discontinuous spatial translation of these same psychological traits. Gradual replacement absolutely does involve discontinuous spatial translation, despite widespread misunderstanding to the contrary. The error is in overlooking the fact that if we replace each neuron with a microscopic prosthetic device that resides nearby, then when the prosthetic eventually disconnects the neuron, leaving it to die, and attaches itself in the neuron's place to take over the neuron's role, the neuron's nonphysical functional operation (and all its metaphysical attachments) will—if one buys the spatial properties of abstractions in the first place—suddenly discontinuously jump over to the prosthetic. The prosthetic does not cohabitate with the neuron after all, as no two physical entities ever can, of course. The prosthetic is generally assumed to reside next to its host neuron, a distance on the order of tens to hundreds of microns. Astoundingly, over the span of a procedure involving a hundred billion neurons, this discontinuous spatial translation of neural function accumulates *tens to hundreds of kilometers* of discontinuous translation of neurological, psychological and metaphysical traits—it's just math, as they say. This is still in reference to a fully-awake, gradual in-place replacement uploading procedure, widely touted by its proponents for its solution to some imposed need to maintain spatial continuity between a brain and its upload infrastructure. Yet it utterly fails to achieve that very goal.

One response to this clarification that gradual in-place replacement uploading also involves spatial discontinuities could be to simply dismiss distances of a few microns as unimportant and somehow capable of preserving identity, but what rationale would justify such an exception? What feature of metaphysics would suggest dismissing small physical distances as metaphysically unimportant but maintain that larger distances have profound effects? As I have argued elsewhere, it is suspiciously convenient to prescribe that we should overlook translations short enough to escape our daily awareness (microns), yet that we should assign great metaphysical significance to translations on the scale of our daily experience, such as from a brain to a computer across a room. Would aliens considerably smaller or larger than us assign such an arbitrary threshold differently?

Another response to in-place uploading's inherent spatial discontinuities is to simply abandon the terminology of, and requirement for, continuous spatial translation entirely and pivot to a completely

different concept: a requirement for piecemeal replacement as opposed to whole parcel replacement, in the vein of the infamous Ship of Theseus or Grandfather's Axe thought experiments. But Koene and I already argued that it is fallacious to believe that gradual replacement is more identity-preserving than scan-and-copy in our paper. With that argument thoroughly addressed (and we believe adequately refuted), the most likely next response is to require the preservation of a dynamic neurological process that is not permitted to cease. Note the significant change in topic here, as we now focus entirely on pure function and no longer on form or matter or space. Perhaps metaphysics should handle function differently than it handles form. This feels like unfair goal-post moving, but I will persist. The claim is that continuous neural activity preserves a continuous cognitive experience along with its continuous SoC and persistent identity. That claim is not too contentious actually, for surely anyone will agree that ceasing neural function also ceases cognitive experiences and metaphysical phenomena. The contention concerns the proper metaphysical interpretation should the physical process be restarted at a later time, either in the original substrate or in a new system that implements the same function. Should the restarted neural process be considered the same process from before, or a new process? Should the cognitive experience and SoC be labeled a continued prior experience, or a brand new one? Should the identity of the person in question be labeled an awakened survivor, or a *de novo* replacement? These questions are where disagreement occurs.

## **Identity relative to visual and spatial continuity**

One point of evidence of our identity's resilience to apparent spatial dislocations comes from our capacity to maintain our identity when experiencing seemingly discontinuous, if perhaps discombobulating, experiences of change in location, such as when awakening after being moved while asleep, or when donning virtual reality goggles and experiencing a spontaneous relocation to a virtual location. Both examples yield the experience of discontinuous relocation, yet are never labeled as identity-erasure risks. Furthermore, we don't *have* to be aware of relocations to maintain our identity since we could be moved while asleep to an identical room. So our evolving mental states need never incorporate the fact of a dormant relocation. Likewise, we are never accused of vanishing forever and suddenly being replaced just because we put on VR goggles and experience illusory discontinuous relocation.

So identity has no dependence on awareness (sleep) or experience (VR) of continuous location. While these examples don't necessarily disprove a dependence of identity on spatial continuity, they do suggest that identity may preserve across discontinuous physical relocations. At the very least, our identity can't possibly *rely* on our awareness of our spatial continuity since our identity is clearly permitted to survive the *experience* of discontinuous spatial translation. Consequently, even if there is an underlying physical-to-metaphysical connection that renders identity dependent on spatial continuity, that connection must operate independent of our awareness. Although our awareness of purported properties of our identity need not be a requirement, we should be careful about the risk of introducing superfluous, unfalsifiable, unverifiable metaphysical traits. If a proposed feature of identity, such as spatial continuity, is claimed to hold true despite a complete lack of any causal effect it could have upon our intimate personal experiences, then how is that claim any different from a total absence of the proposed extraneous feature in the first place? Occam's Razor urges us to jettison claims of this sort in the interests of parsimony. The only stream that *appears* to be relevant to our identity is our phenomenological sense of ourselves as we consume environmental stimuli in combination with our memories (and all related cognitive aspects).

The likely response to our visual experience is that the critical dependency is not on our visual stream but rather on the physical and spatial continuity of our bodies regardless of our awareness of that continuity (sleep) or any illusory experience to the contrary (VR). So an experience of discontinuous location (awakening after being moved or entering VR) is dismissed as noncritical to identity, while our

physical location, which may not have any effect whatsoever on our experience when asleep, as explained above, is somehow deemed critical. It is frustrating that a form of continuity with direct impact on our conscious experience, such as our visual sense of location, is demonstrably irrelevant to our identity, while another form of continuity, such as spatial continuity, which does not necessarily have a causal relation to our experience, and therefore seems definitively superfluous, is frequently assigned grandiose preeminence to our identity. This reasoning seems precisely backwards to me.

The proposal, then, is that identity has nothing to do with visual stimuli (surely to the relief of blind people), but is utterly dependent on the spatial continuity of the body, a continuity of which we need not even be aware. Let us call this theory the *awareness-independent-spatial-continuity-dependence* property (AISCN or "acid" for ease) of metaphysical personal identity. This proposed AISCN feature of identity is then taken to be of paramount importance despite the lack of a tight connection between physical continuity and our conscious visual experience. What is the justification for AISCN? Why not simply discard it?

How should we tie metaphysical identity to AISCN? The examples above suggest that conscious awareness of our spatial continuity is totally unimportant to our identity; our identity survives breakages in that experience, both if we move but have no knowledge of it (sleep) and if we only seem to move but don't (VR). Consequently, it is unclear why AISCN should be a reasonable proposal. AISCN requires assigning an as-yet completely undiscovered physical property of the universe that seemingly ties our brain's location, and its smooth continuous movement, to continuous mental function, and then on to unbroken metaphysical identity, and furthermore in ways seemingly totally disconnected from our experiences and awareness. If one is going to propose that identity has dependencies far beyond our conscious experience, then the nature of such a metaphysical property must be established. What aspect of mental function requires physical continuity? What physical law precludes pausing an information-processing function for a while and resuming it at a later time, or even resuming it in some other substrate in another location? Remember, gradual in-place replacement suffers precisely the same continuity problems. It just does so at the microscopic level of neurons, escaping our ordinary experience, instead of at the level of the entire brain where we can immediately perceive the spatial disconnect involved. It is a mere oversight that this fact of spatial discontinuity in gradual replacement has been completely missed by most people in the past. Yet many people are comfortable assigning gradual replacement an identity preservation status.

I have argued that the dependence of identity on spatial continuity is weaker than generally assumed, but is there any particular reason to favor either position: requiring or dismissing spatial continuity, or is it even odds? Occam's Razor gives us the reason. Extraneous, superfluous features of a model of reality, whose presence, even if true, would have no causal implications and would be empirically undetectable and unfalsifiable, are best jettisoned from our model entirely. At the least, surely the burden of proof lies with the proponent of such acausal superfluous traits.

## **The memory basis for identity**

Consider the following question: How are you confident of your identity relative to past "versions of yourself", such as yourself from:

- five minutes ago.
- before you went to sleep.
- before you fainted.
- before you suffered a concussion, grand mal seizure, stroke, or other neurological trauma.
- before you underwent general anesthesia.

- before you underwent medically induced hypothermia to protect your brain during a cardiac emergency.
- before you suffered rapid frigid drowning in which you fell in a freezing cold lake, drowned, remained submerged for an hour, were taken to the hospital, confirmed to have essentially no pulse and absolutely no brain activity, and then finally revived.
- before undergoing some as-yet-speculative long-term preservation or stasis, such as hibernation or cryonics.

The situations shown above are known as the *bridge problem*: If our identity depends on a continuous SoC, then how does it bridge gaps in that stream? Clearly, the original terminology of requiring preservation of the SoC to preserve identity was imprecise from the beginning, on the realization that even SCPs generally acknowledge that it is their own persistent identity that awakens each morning. Dainton and Bayne solve the bridge problem by disposing of psychological continuity (memories and other cognitive traits) in favor of what they call *phenomenal continuity*, a stream of experience [Dainton & Bayne 2005]. They then solve bridging by naming an identity as that which has the *capability* for continuous experience even if, in actuality, gaps occur. I disagree with aspects of Dainton and Bayne's thesis in that I characterize identity as our experience of our memories and other cognitive aspects in the specious present amalgamated with any new sensory stimuli. They propose hypothetical scenarios in which a SoC is separated from a stream of psychological memories and then they conclude that it is blatantly obvious that identity follows the stream of phenomenal experience. But to me it is equally clear in their scenarios that identity goes with the memories instead! Somehow, they did not foresee such disagreement in their original writing since they present their conclusion as practically beyond dispute. I discussed this phenomenon of philosophers talking past one another in the introduction. Nevertheless, I find their bridging solution to be reasonable: identity consists of a persistent *potential* across a gap even if not fully realized. Where we disagree is the nature of that potential: a stream of conscious experience versus a speciously present experience of memories spanning the gap.

Assuming we continue to pursue the more popular SoC claim, consisting of a never-ending stream of experience, not Dainton and Bayne's recast potential for such experience, then we will see below that such a claim is highly problematic. SCPs usually handle this challenge by leaving the SoC claim behind and moving on to other arguments, such as spatial continuity or continuous neural activity. But then why not open on those terms and abandon the SoC claim once and for all?

One response to the questions shown above is that we exclusively identify ourselves by the phenomenological experience (i.e., experience in the moment) of our memories of our lives. Namely, there is explicitly not any significance to some never-ending SoC, only to our awareness and experience of our memories at any instant in time. Obviously, such awareness and experience can wax and wane without loss of identity, as illustrated by the fact that we generally judge identity to be preserved across gaps (e.g., sleep). There is a similar thought experiment put forward by Philip Gosse in the nineteenth century, and then again by Bertrand Russell in the twentieth, asking how we know God did not create the universe a few seconds ago with our memories artificially created *de novo* (Russell's version doesn't involve God, but the proposal is the same). The answer is that we cannot know that, precisely because our memories are the sole indicator of our confidence in reality, including our confidence in our own identity. It would seem that our momentary memories—our instantaneous experience of ourselves and the world—are the exclusive indicator of our phenomenological consciousness and identity for all intents and purposes. It just isn't clear why we should bother adding any additional features beyond our speciously-present (synchronic) conscious experience of our memories.

Ironically, the patients most susceptible to claims of unpreserved identity are not futuristic science-fiction mind uploads or clones, but rather real-world patients for whom neurological damage has

unwoven the contiguous memories that would underlie their lifelong identity, such as from stroke or dementia, or damage that prevents the formation of new memories, i.e., anterograde amnesia (thereby wrecking the connective experience between the present and past), or damage that obliterates existing memories that would lay the bedrock of their deeper identity, i.e., retrograde amnesia. These are the people, if anyone, best argued as having lost their metaphysical personal identity, despite preserving their original brains and bodies.

It is worth briefly remarking that SCPs must presumably reject cryonics, not necessarily in terms of any practical plausibility (a grave question I don't address in abundance here), but in terms of any hope of metaphysical preservation of identity. If a viable form of cryonic preservation and revival were ever developed, the SoC community should brand revived patients as dead originals replaced with newly invoked identities. After all, cryonic preservation surely satisfies their criteria of a broken SoC since no dynamic neurological process persists during stasis. Yet in virtually every way imaginable, a revived cryonics patient has undergone a physical process and mental experience almost indistinguishable from general anesthesia or medically induced hypothermia. Why should we hold such a strange dissonance between these procedures instead of bringing them into alignment, either by rejecting the identity of millions of actual surgical patients, or by accepting the identity of ostensible cryonics patients? But to accept the identity of revived cryonics patients undermines the claim that never-ending process is important. Dainton and Bayne's solution to cryonics is obvious and reasonable: cryonics patients maintain the capability for consciousness and experience, and in so doing, bridge their identity across stasis. But proponents of a true persistent stream face a conundrum in cryonics.

Certain medical cases actually do appear to involve a reduction in a patient's identity, in that external observers such patients appear altered in personal aspect. It can seem like some or all of the original person has been lost, while at least a little bit of some newly invoked person has taken residence in their place. Examples include strokes, dementia, and some specific examples from the medical literature in which neurological damage radically alters personality, such as the famous Phineas Gage, and a few patients whose brain tumors produce strange personality changes. However, in these real-world cases, no one is ever truly regarded as a doppelgänger. We just don't treat people that way. That said, a committed SCP might argue that identity loss in such cases results from metaphysical damage (whatever that would entail) to their SoC, but it is just as easily argued that any alteration to identity in such cases results from the loss of connective cognitive traits (memories) that would have otherwise tied their phenomenological consciousness in the present to their memories of their past.

## **The neural threshold of the stream of consciousness**

Another common response to the scenarios presented in the list above is that in none of those scenarios does neural activity truly cease. During a fainting spell, or while under general anesthesia, underlying neural activity and associated SoC are argued to have been maintained above some apparent minimum threshold capable of preserving identity. This response comes up repeatedly in debates, so scrutinizing it is central to this article. We are clearly not considering Willam James' original coinage of the SoC, as explained at the beginning, but rather this other concept involving some never-ending dynamic neurological process. This counterclaim that some basal dynamic process is preserved is problematic however, in two different ways.

The first response is to simply understand rapid frigid drowning in greater detail [Hilmo et. al. 2014]. People offering the counterargument of basal neural activity must simply not understand what rapid frigid drowning actually involves, or else they would realize the suspension of disbelief that they are proposing. On one hand, some people actually take their expectation of rapid frigid drowning too far, mistakenly

believing that these medical cases are real-world proofs-of-concept of cryonics, on the assumption that these patients have frozen solid and later revived. That is not true. A human body submerged in freezing water contains too much thermal mass, starting from far too high a temperature of 37C degrees, to freeze solid in an hour. In the most extreme cases to date, rapid frigid drowning victims have survived descents to body temperatures around 13.0 to 13.7C degrees [Gilbert et al. 2000, MSN 2014].

At first glance, the acknowledgement of a warm body might seem like a concession to the claim that neural activity could survive. After all, these patients did not truly undergo cryogenic preservation or later revival. Perhaps cooled to only the balmy temperature of a chilly autumn day, one might initially assume that the brain continues enough interneural signal propagation to keep the hallowed SoC tenuously suspended above the chasm of annihilation. However, this conclusion is not supported by the medical facts. In such patients there truly is no measurable neural activity. Admittedly, at 13C degrees, metabolism still occurs at a molecular level. Nevertheless there is no *neurological* activity. Action potentials do not propagate along axons and dendrites at 13C degrees. Neurotransmitters do not meaningfully traverse synapses since such behavior is causally associated with action potentials, much less do such neurotransmitters trigger new action potentials in downstream neurons. There is no neural information processing, cognitive thought, or conscious experience. To the contrary, the temperature at which the brain effectively goes mute is much warmer, closer to 21C degrees [Lomber et. al. 1999]. This is well established by multiple experiments involving a medical procedure known as *medically induced hypothermia*, in which a patient's temperature is intentionally lowered, often following a cardiac emergency, specifically as a neuroprotective measure to prevent brain atrophy [Hemmen & Lyden 2007, Mizrahi et. al. 1989, Percy et. al. 2009]. The fact that rapid frigid drowning indicates real-world cases in which patients have not been treated as doppelgängers following their ordeals clearly demonstrates that the SoC claim is incompatible with our general concept of reality. Dainton and Bayne are right to solve bridging with an unrealized *potential* that spans the gap as opposed to continuing to insist that at some inscrutable level, no gap occurs in the first place. I just think they built the wrong bridge in their reliance on experience instead of memories.

Yet, admittedly, despite the medical facts of rapid frigid drowning, the SoC argument remains possible in the strictest logical sense, if nevertheless untenable in any reasonable sense. For example, it could be argued—I am unsure how—that the brain of a rapid frigid drowning patient is mysteriously preserving just barely enough neural activity and cognition to string the SoC along on a metaphysical spider's thread. Perhaps our brain scan equipment cannot detect such diminished neural activity at our current technological level. This proposal is all but implausible, but let's briefly explore it anyway.

Let us assume that the proposed SoC does exist and is critical to identity. It must be underlied by some correlate of neurological activity, a point clearly not in dispute since SCPs are arguing that the brain's neurological function was not sufficiently diminished in any of the examples shown on the list. We are then faced with the following remarkable conclusion: there must be some physical threshold of neural activity, with its associated metaphysical threshold of SoC, below which a revived patient should be branded a doppelgänger. It was explained above that existing medical cases of rapid frigid drowning have achieved approximately 13C degrees body temperature. Given that in such cases, it has never been seriously proposed that these patients awaken as identity replacements, we may conclude that 13C degrees has already been deemed metaphysically safe for our identity. Perhaps the temperature at which neural activity becomes so muted that the critical SoC suddenly breaks forever is 12C degrees. This extended argument strains credulity now that it has been demonstrated that real-world 13C degree patients exhibit no neural activity, but I will persevere.

The claim then, is that if we were to revive a patient from 12C degrees (or whatever threshold is ultimately discovered) we should brand them a metaphysical copy. They need not display any outward cognitive deviation to earn this fate. Their memories and personality could survive. Everything about

them could indicate to an observer that they survived, yet the claim is that we should label them as a mental duplicate merely for having dropped below the predetermined 12C degree mark and its apparent consciousness stream threshold.

To be clear, we are not speculating on the possibility that a person undergoes genuine neurological damage at some colder temperature. We are allowed to assume healthy revival because that is precisely the scenario SCPs argue against. They claim that a seemingly cognitively successful, but discontinuously achieved, upload, with all his or her memories and personality intact, is nevertheless a metaphysical identity copy anyway.

The question is, would we ever actually regard a patient in the described way? Given such a medical revival, would any SCP ever actually point at a patient and say, "I saw her medical chart. She was below 12C degrees, so she's a copy." I am not talking about exploiting philosophical abstractions for arbitrary legal advantages. A greedy heir could ask a judge to legally label an upload as a copy to manipulate the laws of inheritance, but that is not a metaphysical question; that is just exploitation of the law for financial gain. The question is whether such an heir would really *believe* such an accusation. Real-world cases of rapid frigid drowning suggest that such beliefs are extremely rare. More to the point, what would a SCP conclude about their own identity were they the patient? Would they dismiss their own conviction about who they feel they are? What if they didn't know the facts of their circumstances and therefore couldn't formulate their identity on knowledge of how they came to be (this comes up again below)?

## Experimental confirmation of the stream of consciousness

There is a remaining question: How will we ever determine the qualifying threshold of neural activity that indicates the invisible threshold of the SoC breakage toward the ultimate threshold of identity-loss-with-replacement-invocation? If we could do this, we could then use a neural activity monitor to assign identity judgments and labels to patients. This car accident victim is an original, but that flu patient is a copy, according to their brain scans. In theory, we should be able to determine the neural threshold of the SoC by associating it with one of two possible correlates:

1. If we could *detect the proposed SoC* with some metaphysical SoC machine, and detect any breakage in that stream, then we could confirm a person as a metaphysical original or copy and determine the associated neural activity threshold from a brain scan.
2. Alternatively, if we could *detect that someone is a metaphysical copy* via some sort of psychological test, then we would know there is a consciousness stream breakage somewhere in the person's brain scan data history and we could take the lowest recorded neural activity as our threshold.

It is important to note that neither of these proposed approaches is even theoretically possible within physics (for the first approach above) or existing psychology (for the second approach). We have no concept of a machine that could, with regard to option one above, detect the maintenance or breakage of the hypothesized SoC (or measure its metaphysical strength, whatever that might mean), if for no better reason than that such a stream's very existence remains entirely speculative. Likewise, with regard to option two above, since the basis of the debate is that a person need not exhibit any psychological indications of their identity copy status, it is essentially defined away that any psychological test could ever reveal a person as a copy or confirm them as an original. Consequently, this entire task is totally unapproachable by science since it concerns exclusively metaphysical features which, by the SCPs' own description, are not physically discernible. This is the very definition of unfalsifiability.

## Practical application of stream of consciousness identity

Assume that the threshold of neural activity associated with an identity status flip from preservation to loss-with-replacement is determined. Given that neural activity is easily correlated with temperature, we could conceivably use temperature to make identity determinations and judgments. Hospitals could include a stream-of-consciousness-verification test when unconscious patients are brought into the ER. Let's say it has been determined that 12C degrees is the critical threshold in question. Two rapid frigid drowning patients arrive at the hospital, Alice and Bob. Alice is revived, but from below 12C degrees. She is branded a copy for the rest of her life. Alice herself, being a strong believer in the SoC claim, accepts her fate as a copy and folds this realization into her sense of her own identity, believing that her memories from before the accident are actually fallaciously those of a completely different person and that she was born when she awoke after the accident. Bob, however, arrived at the hospital with a temperature of 13C degrees, much like real-world patients for whom no fictional speculation is required. Bob is clearly granted the status of preserved identity, consistent with contemporary medical cases. Alice and Bob proceed to live their lives for many decades, interacting with numerous people who regard Alice as a copy and Bob as a preserved identity. Only toward the end of their lives does someone investigate their old medical records and discover to their amazement that both measurements were wrong! Alice never actually descended to 12C degrees in the first place. There was an error in the original measurement and she only descended to 13C degrees. Meanwhile, Bob's temperature measurement was also wrong. He dipped to 12C degrees that fateful day. So Alice was never a copy in the first place, yet everyone believed she was, including herself! And Bob was a copy the whole time, yet not one person ever realized it or regarded him as a copy, including himself.

I see two ways to handle the scenario above, and neither bodes well for the SoC claim. (1) We could still hold the SoC claim, believing that it indicates our true identity, or (2) we could conclude that the SoC is not necessarily real in the first place, or at least that a gap in that stream has no implications for our identity. Even if we choose the first option, we are left with the following troubling conclusion: as illustrated in the example of Alice and Bob, the SoC and its connection to identity, even if real in some abstract metaphysical sense, clearly has absolutely no *objective* impact on reality; it is utterly acausal. It is a completely arbitrary, unfalsifiable, and superfluous metaphysical property injected from the outside. It also has no external mental effects on our state of mind, only the circular effects we may *choose* to take aboard by believing the claim to begin with, mental effects just as easily obviated by simply choosing to break the circular reasoning and not believe the claim at the outset. If the claim only has any meaning by our choice to explicitly assign that meaning, then we can just as easily choose otherwise instead. I urge the second option, that the SoC claim is invalid to begin with.

## Conclusion

Note the gravity of the consequences if it turns out that SCPs are in error on the metaphysical question of identity. They reject the philosophical preservation of identity by the scan-and-copy procedure, which although currently theoretical, is likely to become technically feasible within some foreseeable future. At the same time, they accept the gradual in-place replacement procedure, which remains so technologically speculative as to be essentially impossible for all practical purposes. If gradual in-place replacement is even possible, it may take centuries longer to develop than scan-and-copy. Rejecting scan-and-copy destines anyone living during that interim period to death. Not to take such matters unnecessarily seriously, but misunderstanding this issue is a matter of life and death, other daily practical concerns notwithstanding [Hayworth 2010].

It is my position that we should reject the SoC claim and any requirement such a purported metaphysical property would impose on mind uploading procedures. At the least, the onus rests on this claim's proponents to prove such a stream exists and to defend the claim that it would not survive certain mind uploading procedures.

Lastly, there is the concern that this is all whimsically unimportant, or worse, an obtuse disregard for more prosaic societal concerns. Some people may find debates of this sort to be pedantic and even snobbish, given the justified concern that advanced futuristic technologies are likely to benefit wealthy elites long before they trickle down to the masses. Worse, some people may expect that such technologies are likely impossible and that such metaphysical navelgazing is an ivory tower distraction in a world of real problems and challenges. To that reaction I say the importance is not necessarily in determining the prospects of technological and medical marvels that reside far in the future, if ever. The more relevant issue, and the reason I have committed so much of my life to contemplating and writing about these questions, is that we profoundly desire the most accurate model possible of reality and understanding of the human condition. Ultimately, we want to understand ourselves as conscious beings in the universe and to understand the nature of our existence. That is the real issue here, at least for me.

## About the author

Keith Wiley is on the board of Carboncopies.org and is a fellow with The Brain Preservation Foundation. He holds a PhD in computer science from the University of New Mexico and works as a data scientist in Seattle, Washington. His book, *A Taxonomy and Metaphysics of Mind-Uploading*, is available on Amazon (<https://www.amazon.com/dp/0692279849>). His other writings, interviews, and videos about mind uploading are available on his website at <http://keithwiley.com> and elsewhere on the web.

## References

Cerullo, M. A. (Feb 2015) Uploading and branching identity, *Minds and Machines*, Springer Netherlands, **25**(1), pp. 17–36. doi: 10.1007/s1102.

<http://link.springer.com/article/10.1007%2Fs11023-014-9352-8>

Corabi, J. & Schneider, S. (2012) The metaphysics of uploading, *Journal of Consciousness Studies* (*JoCS*), **19**(7–8), pp. 26–44.

Dainton B & Bayne T. (2005) Consciousness as a guide to personal persistence, *Australasian Journal of Philosophy*, **83**(4), pp. 549–571.

Gilbert M., Busund R., Skagseth A., Nilsen P. & Solbø J. (2000) Resuscitation from accidental hypothermia of 13.7°C with circulatory arrest, *The Lancet*, **355**(9201), pp. 375–376. doi:10.1016/S0140-6736(00)01021-7.

Hayworth, K. (2010) Killed by Bad Philosophy, The Brain Preservation Foundation. <http://www.brainpreservation.org/content-2/killed-bad-philosophy/>

Hemmen, T. & Lyden P. (2007) Induced hypothermia for acute stroke, *Stroke*, **38**(2), pp. 794–799. doi: 10.1161/01.STR.0000247920.15708.f8

Hilmo J., Naesheim, T. & Gilbert M. (2014) “Nobody is dead until warm and dead”: Prolonged resuscitation is warranted in arrested hypothermic victims also in remote areas – A retrospective study from northern Norway, *Resuscitation*, **85**(9), pp. 1204–1211. doi: 10.1016/j.resuscitation.2014.04.029

James, W. (1890) *The Principles of Psychology*, New York, NY: Henry Holt.

Kuhn T. (1962) *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.

McIntyre R. & Fahy G. M. (2015) Aldehyde-stabilized cryopreservation, *Cryobiology*, **71**, pp. 448–458.

Lomber, S, Payne B. & Horel J. (1999) The cryoloop: an adaptable reversible cooling deactivation method for behavioral or electrophysiological assessment of neural function. *Journal of Neuroscience Methods*, **86**(2), pp. 179–94.

Mizrahi, E., Patel V., Crawford E., Coselli J. & Hess K. (1989) Hypothermic-induced electrocerebral silence, prolonged circulatory arrest, and cerebral protection during cardiovascular surgery. *Electroencephalography and clinical neurophysiology*.

MSN (2014), Girl survives 13 degree body temperature, MSN News. <https://www.msn.com/en-gb/news/other/girl-survives-13-degree-body-temperature/ar-AAmSEW>

Percy, A., Widman S., Rizzo J., Tranquilli M. & Eleftheriades J. (2009) Deep hypothermic circulatory arrest in patients with high cognitive needs: full preservation of cognitive abilities, *The Annals of thoracic surgery*, **87**(1), pp.117–123. doi: 10.1016/j.athoracsur.2008.10.025

Popper, K. (1962) *Conjectures and Refutations, The Growth of Scientific Knowledge*, Basic Books, New York.

Wiley, K. B. (2014) *A Taxonomy and Metaphysics of Mind-Uploading*, Humanity+ Press and Alautun Press. <https://www.amazon.com/dp/0692279849>

Wiley, K. B. & Koene, R. A. (2016) The fallacy of favoring gradual replacement mind uploading over scan-and-copy, *Journal of Consciousness Studies (JoCS)*, **23**(3–4), pp. 212–235. <https://arxiv.org/abs/1504.06320>

Wiley R. H. (2015) *Noise Matters: Evolution of Communication in Noise*, Harvard University Press, Cambridge.