Identity Theft: A Thought Experiment on the Fragility of Identity

David Menčik
University of Novi Sad, Serbia
E-mail address: davidmencik@gmail.com
ORCID ID: https://orcid.org/0000-0003-3207-4795

Abstract
This paper intends to discuss some aspects of what we conceive as personal identity: what it consists in, as well as its alleged fragility. First I will try to justify the methodology used in this paper, that is, the use of allegories in ontological debates, especially in the form of thought experiments and science fiction movies. Then I will introduce an original thought experiment I call “Who am I actually?,” one that was coined with the intent to shed light on several aspects of the issue under examination, that is, the fragility of personal identity. Then I will move on to Christopher Nolan’s film The Prestige, as well as to Derek Parfit’s ‘divided minds’ thought experiment, to further discuss the fragility of personal identity; next to identity theft, the prospect of duplication is also intriguing, especially with regard to the psychological impact this might have on both the prototype and the duplicate. I will conclude with the view that spatial and temporal proximity or coexistence, especially when paired with awareness on behalf of the duplicates, would expectedly result in the infringement of the psychological continuity of one’s identity.

Key-words: social identity; physical identity; psychological identity; philosophy and fiction; divided minds; The Prestige; spatial and temporal coexistence

I. Thought experiments, movies, and the debate on personal identity: Justifying the connection

Ever since the classical era there has been “a long-standing antagonism between poetry and philosophy.”1 Plato’s idea of the philosopher is that of a person who has somehow managed to exit the cave, has seen the truth, and returns with the obligation to reveal to others the truth he has become aware of. In Plato’s view the philosopher’s task is to present things the way they truly are and persuade others by means of sound argumentation. Contrary to philosophy, poetry is only entitled to the

subjective interpretation of the truth. Poetry and painting are seen by Plato as two forms of monoperspective representation prone to presenting things not as they truly are, but as they might, or might not be.\(^2\)

This ‘Platonic’ antagonism between philosophy and art in general is still a valid standpoint. However, arguments in favor of this alleged dichotomy between philosophy and literature are being challenged by contemporary thinkers, and several scholars reject the thesis that such an opposition exists indeed. The basic arguments of the counter-dichotomy (or, counter-discontinuity) thesis are that:

i. At least some works of literature present genuine and maybe even novel philosophical arguments that are of the same quality as the best of any bona fide philosophy.\(^3\)

ii. We have no reason to consider literature and philosophy to be dramatically different: (as they) both confront us with nonactual situations.\(^4\)

Based on the counter-discontinuity approach one could make a correlation between philosophy and fiction, or, and this is exactly where this paper will focus, between thought experiments and science-fiction movies. Both thought experiments and sci-fi movies are indeed imaginative ways of illustrating and comprehending a problem. Damir Smiljanić suggests that especially issues related to “possible worlds, problems of identity and ethical dilemmas”\(^5\) could be presented equally well by means of thought experiments and/or sci-fi books and movies, because both combine philosophical argumentation and narration.\(^6\) One could come up with an argument against the discontinuity thesis by providing correlating examples from thought experiments and sci-fi movies that suggest possible answers to issues related to the fragility of identity. The fact that movies about identity-related issues are often filmed in

\(^2\) The reference is to Book II of *The Republic*: “When a storyteller gives us the wrong impression of the nature of gods and heroes. It’s like an artist producing the pictures which don’t look like things he was trying to draw.” Plato, *The Republic*, 377c.


\(^4\) Ibid., 351.


\(^6\) The view that thought experiments can be seen as a composition of argumentation and narration is similar to the one reached by Sören Häggqvist: “Thought experiments are not identical to arguments, they have to be seen as intimately connected to certain arguments.” See Sören Häggqvist, “A Model for Thought Experiments,” *Canadian Journal of Philosophy* 39, no. 1 (2009): 57.
such a way as to force the spectators to participate as puzzle-solvers, makes these movies even more philosophical in nature.\textsuperscript{7}

Furthermore, identity-related movies can become an inspiration for philosophers to create engaging thought experiments.\textsuperscript{8} No less significantly the storyline of a sci-fi book or movie can be placed in a different context, where the correlation between the problem presented in the movie is interwinned with a similar problem presented in the thought experiment, or another movie. Last but not least, the connection between thought experiments and movies is also valid due to that thought experiments are usually designed in such a way, as to highlight the most important part of the issue that is under examination, exactly as movies are structured around the most important scene, that is sometimes repeated during the movie.

II. Who am I actually? A thought experiment on the possibility of social, physical and psychological identity change

Let’s discuss the following – imaginary, though logically permissible – case.\textsuperscript{9} Michael and Rodney are monozygotic twins, but their lives are completely different: Michael is successful, affluent, and has a family of his own; currently he is the CEO of a high-profile company. On the other hand, life hasn’t been that kind to Rodney: he lives alone and is currently unemployed. What is most important for this thought experiment, is that Rodney and Michael look identical: \textit{Their only physical dissimilarity is a tiny birthmark: a faint scar on one of Rodney’s left foot toes (1).}\textsuperscript{10}

On their birthday night Michael and Rodney go out to celebrate; after several hours of consuming large quantities of alcohol, Michael and Rodney have to decide whether they should drive back home, or take a taxi; Michael insists that he was sober enough to drive, and convinces Rodney to get into the car. This reckless decision of theirs proved fatal: a few minutes later their car crashed with another. The driver in the other car was killed on the spot; Michael got severely injured and was out cold, while Rodney only suffered


\textsuperscript{8} For example, the thought experiment I will present later on has been partly inspired by Damir Smiljanić’s lectures on the Philosophy of Mind, during which he discussed extensively John Woo’s film \textit{Face Off}, and partly by Robert Louis Stevenson’s novel \textit{The Strange Case of Dr Jekyll and Mr. Hyde}.

\textsuperscript{9} Thought experiments differ from scientific ones; they allow us to test our assumptions in non-factual situations. My thought experiment consists in a possible – yet, not real – course of events.

\textsuperscript{10} The key steps of the thought experiment are highlighted in italics, and are also numbered.
minor injuries. Michael’s injuries looked fatal to Rodney; convinced that his brother was going to die anyway, Rodney decides to steal Michael’s identity and live his brother’s – much more lucrative – life. Rodney replicates his birthmark on Michael’s toe, an identification mark that only Rodney had prior to this moment (2). Rodney switches documents and wears his brother’s wedding ring on his finger (3).

As soon as the ambulance arrives Michael is taken to the hospital, and Rodney tells the police that he, Rodney,\(^{11}\) was the driver. Afterwards, Rodney and Marie (Michael’s wife to whom Rodney present himself as his brother, Michael) visit Michael in the hospital; Michael’s condition is critical, so the last think one would think of is to check for the faint scar on his toe. Michael has Rodney’s identity card on him, and no wedding ring on his hand. Thus, everyone is convinced that it is Rodney the person who lies in bed heavily injured. This is the moment when the first identity change occurs; it is a change of social identity.\(^{12}\) Rodney presents himself as Michael, and injured Michael is presented as Rodney (4).

Michael is in a comatose state; Rodney is worried that his tiny little birthmark could compromise the identity theft; therefore, he undergoes plastic surgery to eliminate his scar. Now only one person has the scar, Michael. Scar: Rodney’s body; No scar: Michael’s body. When the scar is removed, Rodney’s body looks as if it is Michael’s, and Michael’s as if it is Rodney’s. This establishes a change in pseudo-physical identity (5).\(^{13}\)

After a couple of months Marie has already noticed that her husband’s behavior (who in fact is Rodney) is not like it used to be. She suspects that the person who presents himself as Michael is actually Rodney. Marie decides to visit the comatose twin in the hospital, so as to check whether the person who is supposed to be Rodney has the tiny birthmark that makes the two brothers discernible. After she sees the scar she rests assured that the comatose patient is indeed Rodney; her line of reasoning is outlined as in premise (5): Scar: Rodney; No scar: Michael.\(^{14}\)

---

\(^{11}\) A man with Rodney’s documents, that is, Michael.

\(^{12}\) The term social identity is not easy to define. The closest to a definition could be this: social identity is who one is in relation to others, or who one is presumed to be.

\(^{13}\) I use the greek word word pseudo as a prefix, so as to highlight the difference between pretending to be someone, resembling someone (physically), and being someone. For example, Rodney pseudo-physically is Michael, because he has changed his appearance so as to look like Michael. If, for example, the twins could exchange bodies or consciousness after the accident in some magical way, we would be justified to talk about (real) physical change and its implications.

\(^{14}\) Marie’s logic is flawed. According to the laws of deductive reasoning what she would only be justified to conclude would be that: Scar: Rodney; No scar: Not Rodney (modus tolendo tollens).
This instance is of key significance for the thought experiment, since it is at this point that the connection between social and physical identity is established. The very moment Marie sees the scar, she becomes absolutely convinced that the person lying in the hospital’s bed is Rodney, and that the person who lives with her as her husband is indeed Michael (6).

While the real Michael is in a coma, his brother does his best to destroy everything that Michael has created. One day, though, a miracle happens: Michael unexpectedly wakes up. Rodney is the first to be informed; he rushes to the hospital, tells Michael that he (Rodney) had stolen Michael’s identity, and proposes a trade off; since his trial for manslaughter after the accident was still in progress, and the jury would probably sentence in several years of imprisonment the drunk driver, that is, Michael (now he has woken up), Michael and Rodney both agree to switch identities once more. Rodney will take Michael’s place and, if needed, do time instead of his brother; Michael, in turn, will compensate Rodney for an amount of one million dollars that will be transferred to Rodney’s account. Michael will be free to go back to his wife and do his best to save his company.

It is important to note that this change of identity is only a social one; Rodney never tells Michael that Rodney has no scar, while Michael has. From this moment on Michael’s social identity is ‘Michael,’ even though his pseudo-physical identity is ‘Rodney,’ and vice versa in the case of Rodney (7).

When Rodney goes to the bank to open the bank account where Michael would transfer the agreed amount, he decides to use his fingerprint as an identification; even identical twins have different fingerprints.

Rodney provides a physical proof for his identity, that confirms his social identity as Rodney. This adds a new dimension to the thought experiment. A physical identification mark that cannot be duplicated would prevent Michael from tricking Rodney and stealing his money (8).

This story could have a happy ending if Rodney hadn’t been killed by his inmates shortly after he was found guilty by the jury and incarcerated. The tables were turned when Marie noticed the scar on Michael’s left foot. It was the same scar she had seen on the hospitalized brother’s toe in the hospital, that lead her to the conclusion that the person lying in bed was Rodney as shown in (6), though in fact it was Michael. Michael tells Marie that Rodney bears exactly the same scar; but when Rodney’s corpse is being checked, Michael is shocked to see that there is no such scar on Rodney’s leg. Marie is now pretty much sure that the person she lives with and presents himself as Michael is in fact Rodney who has stolen Michael’s identity, and that the one who died in prison is actually Michael.

Rodney dies in prison; Marie notices the scar on Michael’s leg; Rodney’s
corpse has no scar. Marie concludes that Rodney, the twin with the scar, stole Michael’s identity; Michael, the twin without the scar, has died in prison (9).

Because of this totally unanticipated course of events, Michael loses everything. When he goes to the bank, he is denied access to the account due to the mismatch between his and Rodney’s fingerprint. For all intents and purposes Michael is now socially Rodney because everybody thinks he is Rodney. At the same time, as far as the bank is concerned, he can’t be Rodney, since his fingerprint doesn’t match. Pseudo-physically Michael should be Rodney if the scar was to be the only determinant, but physically he is still Michael. Michael doesn’t know who he is anymore; everything on him tells others that he is Rodney, but he knows he is not. In a moment of despair Michael loses it and kills himself. His last words left in a note were: “Who am I actually?”

Michael’s fingerprint is different; socially he is both Rodney and not Rodney; pseudo-physically he is Rodney; Michael doesn’t know who he is and kills himself (10).

III. Implications of the “Who am I actually” thought experiment and correlations with Christopher Nolan’s *The Prestige* and Derek Parfit’s ‘Divided Minds’

a. Determinism and social identity changes

This thought experiment includes four changes in social identity. An issue that needs to be addressed is whether these changes occurred due to the free will of the persons involved, or were determined by factors external to the agents. It is obvious that the first change in both brothers’ social identity is voluntary for Rodney, and involuntarily for Michael (3, 4). But still, even in the case of Michael, it cannot be deemed pre-determined. The identity exchange in (7), when the brothers decided to switch back again, has been voluntarily for both. This identity change is partially pre-determined and partially not, since it was step (3) that led to it, but the brothers didn’t need to change their social identities back. The third major social identity change happens in (9), when Michael is socially restored as Rodney. This change is completely involuntary since it is determined by steps (1, 2, 3, 5, 6, 7), that is, all the steps that precede it, and include previous social identity changes, as well as pseudo-physical identity changes. Another important factor as far as the pre-determined character of this identity change is concerned, is the fact that Marie reaches the wrong conclusion regarding Michael’s physical identity while Michael was in a coma. The final social identity (10) is also involuntary and determined by (8), and indirectly determined by all previous steps. The fact that, as it is shown in the thought experiment, it is not
always us who decide whether our social identity will stay the same or not, leads to the question on why is social identity often as fragile as it has shown to be in the case of Michael.

b. The fragility of social identity, Michael’s case, and the ‘canary switch’

The thought experiment I presented above was intended to show that social identity can easily be altered. It may even seem that social identity is just an illusion. It is easy to fool other people in believing that one is the person one claims to be. In Michael’s case the fragility of social identity is shown in step (3), when his identity is stolen by Rodney by simple actions like exchanging documents and stealing a wedding ring. Furthermore, Michael’s case shows that one can easily be mistaken for the person one is not (6, 9), and also that it may be not as easy to provide a solid proof of one’s social identity in the case an identity theft occurs (10).

The ‘canary switch’ scene in Christopher Nolan’s film The Prestige presents masterfully the fragility of social identity. In the scene a canary is put in a cage on a table, and the cage is covered with a magician’s scarf. The magician smashes the cage with his hand; then he removes the scarf and the cage as well as the canary have just vanished into thin air. Right after, the magician conjures the canary up safe in the scarf. The trick is performed twice during the film, and both times Nolan focuses on the impact it has on the spectators – it is awe and horror. When the full version of the trick is exhibited, the truth is revealed: the table has a secret trap hatch, and right below the hatch is attached a compartment for the cage. The canary is being smashed indeed, and the trick lies in that it is artfully replaced by a ‘twin’ canary. What connects my thought experiment to Nolan’s canary switch is the illusion of empirical singularity.

An illusion involving the use of doubles functions through the dissimulation of the plurality of an object whereby the pledged

15 One’s social identity is not one’s possession, nor an intrinsic attribute of one’s essence. It is, after all, the only kind of identity that can be stolen, as it is shown in my thought experiment. For example, Rodney was able to steal Michael’s social identity, just because he needed to convince other people that he is Michael. Fooling others was easy, since Michael and Rodney looked almost the same. On the other hand, Rodney could have never stolen Michael’s psychological identity, because this kind of identity is inherent in one’s essence as a part of one’s consciousness.

16 The feeling of horror emerges when one realizes what the trick consists in, as it happens when Nolan shows a boy having realized that the canary presented at the end of the trick is not the initial one, and asking: “But where’s his brother?” Christopher Nolan, and Jonathan Nolan, The Prestige: Screenplay (London: Faber & Faber, 2006), 19.
object does disappear or even die, but is immediately replaced by something else which looks exactly the same as the thing which was made to disappear.\textsuperscript{17}

Though my scenario doesn’t involve any disappearance trick, it nevertheless addresses the issue of the logical association between uniqueness (or, singularity) and duality, as the identity of both Rodney and Michael becomes dual at several instances. Marie, for example, believes that the birthmark on Rodney’s toe is unique, and that it is therefore a proof that the person that bears it is Rodney. Rodney’s identity can also be understood as dual because he is both Rodney (physically and psychologically) and Michael (pseudo-physically and occasionally socially) at the same time. Rodney’s duality consists in the fact that he is both himself as well as an impersonation of Michael, and the same applies to Michael, who is also an impersonation of Rodney (pseudo-physically and socially).

In my view all four types of personal identity I outlined in my thought experiment are interrelated.\textsuperscript{18} The ‘canary switch’ scene in Christopher Nolan’s film seems also capable of providing support to arguments in favor of the fragility of social identity. However, the fragility-of-identity hypothesis seems to be in need of further support; such support, I believe, could be sought in Derek Parfit’s ‘divided minds’ thought experiment.

c. Psychological and physical identity: The originality of a replicated being

The problem of creating an identical copy of something that is presumed to be unique has fueled a persistent ontological debate that starts with Plato’s discussion of \textit{mimesis} and is present all the way through up to the critique of the ‘cultrual industry’ by Horkheimer. In “The Work of Art in the Age of Mechanical Reproduction” Walter Benjamin describes the mechanism of reproduction as follows:

\begin{quote}
The technique of reproduction detaches the reproduced object from the domain of tradition. By making many reproductions it substitutes a plurality of copies for a unique existence.\textsuperscript{19}
\end{quote}


\textsuperscript{18} Pseudo-physical identity and social identity are similarly structured, and are equally external to one’s essence. I will argue that we could speak of four identities, provided that we keep in mind that the pseudo-physical identity doesn’t exist per se; it requires a connection between social and physical identity. This identity emerges during our interaction with others (social identity) on the basis of how we look like (physical identity). This is why we may change our identity to look like someone else (pseudo-physical identity), so that other people will think that we are another person.

\textsuperscript{19} Walter Benjamin, “The Work of Art in the Age of Mechanical Reproduction,” in \textit{Illuminations},
In their *Dialectic of Enlightenment* Theodor Adorno and Max Horkheimer argue that the cultural industry has set out to reproduce everything that exists in the world: real-world objects are replicated by literature and the movie industry, and this results in the alienation from one’s self, others, and real world problems. The distinction between the world of reality and the world of fiction is lost; people’s attitude towards the world can be altered by regulating what they are furnished during their time of leisure, therefore controlling the media means enforcing conformity and manipulation. In their works Adorno, Horkheimer, and Benjamin have discussed the devaluation of actuality in a world that relies on omnipresent reproduction. My focus here is not on political theory, of course; the issue that concerns originality and the way it is related to value, however, is definitely a common thread.

The question is whether creating identical copies of existing persons would compromise those persons’ uniqueness or not. Leaving aside the various possible bioethical dilemmas that arise from human cloning, and limiting the focus on the issue of identity, there are two questions that need to be addressed:

a. Could my replica be physically identical to me?
b. Would the existence of a living replica of mine compromise my psychological identity?

The answer to the first question doesn’t seem to be that complicated. If we had the power to create a machine that would produce replicas, like the one Tesla created for Angier in Nolan’s *The Prestige*, I could be physically identical to my copy if 100% of my matter was used to give existence to my replica without destroying me at the same time. However, it is important to note that even mere physical similarity, or should we say pseudo-physical identity between replicas and originals, could infringe psychological identity. We can see a reaction towards this infringement in Angier’s expression after his first replica was created: Angier is stunned and horrified at the same time. Angier seems to be afraid that his own identity is being compromised by the

---


mere existence of a replica of his, even though Angier knows that it is only about a duplicate, and not the ‘real him;’ this is why he eventually decides to annihilate it. It seems that it is almost impossible to come to terms with the fact that I am staring at something that is and is not me at the same time.

Parfit’s discussion is much more complex; in his ‘divided minds’ thought experiment Parfit discusses an imaginary situation in the context of which he is being ‘teletransported’ on planet Mars. During the process he loses consciousness and his body on Earth is being destroyed as scheduled, while at the same time a replicator on Mars produces a body that is absolutely identical with the one on Earth; his consciousness is also replicated: this new self has exactly the same knowledge, experiences and memories as the prototype.22 Both identical beings (this on Earth, and that on Mars) share a common physical and psychological identity; the distance between them, however, made possible due to the duplicate’s ‘teletrasportation’ and the destruction of the prototype, allow for the psychological impression that the replica is indeed a unique self. In this part the clone in Parfit’s thought experiment doesn’t share Angier’s frustration due to his replication.

In the second part of the thought experiment Parfit enters the machine, but doesn’t get teletransportated; he is being told, instead, that “the new scanner records your blueprint without destroying your brain and body.”23 An interesting question is whether he would consent to this if he was in advance informed that both versions of his self would coexist, albeit in different places. It is inconceivable for one to be in two places at the same time. “Wait a minute. If I am here, I can’t also be on Mars.”24

Shortly after, the prototype on Earth is told that his cardiac system has been damaged due to the replication process, and that he is going to die, while the replica on Mars will keep on living. When the original and the replica communicate, they are both convinced by their conversation that, even though physically identical, psychologically they are not the same person. “Call this the Branch-Line Case [...] though he is exactly like me, he is one person, and I am another.”25

IV. A possible conclusion on the fragility of identity

In the first part of my article I set out to provide reasons for the methodology I used – both as far as my research, as well as the presentation of my results

23 Ibid., 198.
24 Ibid.
25 Ibid., 201.
are concerned. Even though the paper is quite short, and the focus has necessarily been narrowed down to only a few relevant cases, I believe that the conclusions I have reached and I am about to present could be of broader applicability and validity, even though just as hypothetical and speculative. Reaching speculative conclusions goes hand in hand with thought experiments, since they are just allegories intended not to prove, but to showcase. On the other hand, as far as identity issues are concerned, thought experiments is the best thing we can afford, since it is unethical to examine the fragility of personal identity by performing experiments on human beings.

Based on the discussion in this paper, one could reach two distinct conclusions. The first is that our social identity is not intrinsic to our essence; it almost sounds improper to speak of an identity at all, since social identity is entirely dependent on the objectification of one's own self by others. Perhaps it would be much more appropriate to understand one's social identity as a means of interaction. If persons didn't interact, there would be no such identity attached to one's self. We could imagine a situation in which a person, let's say, who suffers from total amnesia is transferred in the wilderness. Living in the absence of any other person, this person is socially 'no-one,' as nobody else interacts with him. Next to this, as it has already been shown previously in this paper, one more argument in favor of the non-actuality of social identity is that it can easily be abolished, or transferred to another person.

The second conclusion would be that one's physical identity is interrelated with one's social and psychological identity. For example, similarity in physical appearance, or what I call pseudo-physical identity, may result in the social identification of two persons, as it is shown in the case of Rodney with regard to his scar, and in the 'canary switch' scene in Cristopher Nolan's The Prestige. Furthermore, as Parfit's thought experiment suggests clear no less than Angier's reaction to his replica, the psychological continuity of one's identity would be infringed if a. the prototype and the replica coexisted in spatial and temporal proximity, b. either of the two was aware of the existance of the other.

**Acknowledgments**

As the twin who survived, I didn't have the opportunity to experience how it is like to be a twin. It is probably due to this that issues as such attract my interest that much. But turning inclination into philosophical articulation is something that always requires the aid of mentors and collaborators. Therefore, I am indebted to Professor Damir Smiljanić for providing the initial and absolutely necessary inspiration and motivation for this paper. I am
also in Dr Goran Rujević’s debt for being a critical reader and an extremely insightful commentator. I also need to thank Milena Stefanović who helped me discover my special interest in science fiction and encouraged me to make use of it in my philosophical inquiries. I am in Ruel Mannette’s debt for providing insightful remarks and suggestions. Finally, I need to express my gratitude to both anonymous reviewers to whom my paper was assigned by the Editors of Conatus for their suggestions.

References


