Madness, Marginalization, & Memory

The Relational and Personal Harms of Feeling "Stuck in Time"

Emily Kate Walsh

Department of Philosophy,

McGill University, Montreal

A thesis submitted to McGill University in partial fulfilment

of the requirements of the Degree of Doctor of Philosophy in June 2022.

In memory of Doreen Griffiths

“The very best relationship has

A gardener and a flower.

The gardener nurtures,

And the flower blooms.”

Dedicated to Alba Fiamma Perego Aliffi

« Non sta scritto da nessuna parte che non ce la puoi fare »

& Sophie Osiecki

“Not all those who wander are lost.”

Table Of Contents for the Thesis

Abstracts

English Abstract

Little attention is given in psychiatry to the temporal aspect of psychiatric illness. My thesis argues that without the ability to mentally time travel (MTT), to direct your future self and reflect on your past self in the way that humans are designed to, mental distress follows shortly after. This thesis explores how being stuck in time, where one feels trapped in a traumatic memory repertoire, a set of circumstances, or a set of beliefs, prevents an individual from being able to MTT. It argues that this feeling of being stuck in time can result from oppressive social structures, psychological conditions, and a combination of both. This thesis thus aims to flesh out the connections between madness, marginalization, and MTT. My thesis begins by unpacking the psychological threats which can occur to one's ability to MTT. I show that whilst MTT is currently theorized as an individualistic and rationalistic capacity; trauma shows us that there are more cognitive, affective, and relational elements than the current literature acknowledges. After exploring this figurative sense of being stuck in time, in which one cannot psychologically move past certain experiences, the following two chapters explore the social sense of being stuck in time through the cases of racialization and dementia patients. Given the relational nature of the self – that the self is constituted by our relations with others – there are degrees to which individuals can be hindered in the task of developing a coherent sense of self. Indeed, when one is socially or mentally struggling, people can be 'held in place' by others to preserve a sense of self. This holding can be supportive, but one central concern in my thesis is that it can also be oppressive. My thesis ends by showing that whilst feeling stuck in time can cause several personal and relational harms, these harms can be overcome by our connections with others in our social circle through interpersonal trust.

French Abstract

Peu d'attention est accordée en psychiatrie à l'aspect temporel de la maladie psychiatrique. Ma thèse soutient que sans la capacité de voyager mentalement dans le temps (MTT), de diriger votre moi futur et de réfléchir sur votre moi passé de la manière dont les humains sont conçus, la détresse mentale suit peu de temps après. Cette thèse explore comment le fait d'être bloqué dans le temps, où l'on se sent piégé dans un répertoire de mémoire traumatique, un ensemble de circonstances ou un ensemble de croyances, empêche un individu de pouvoir faire du MTT. Je soutient que ce sentiment d'être coincé dans le temps peut résulter de structures sociales oppressives, de conditions psychologiques et d'une combinaison des deux. Cette thèse vise ainsi à étoffer les liens entre folie, marginalisation et MTT. Ma thèse commence par déballer les menaces psychologiques qui peuvent survenir à sa capacité à MTT. Je montre que si le MTT est actuellement théorisé comme une capacité individualiste et rationaliste ; le traumatisme nous montre qu'il y a plus d'éléments cognitifs, affectifs et relationnels que ne le reconnaît la littérature actuelle. Après avoir exploré ce sentiment figuratif d'être bloqué dans le temps, dans lequel on ne peut psychologiquement dépasser certaines expériences, les deux chapitres suivants explorent le sentiment social d'être bloqué dans le temps à travers les cas de racialisation et de patients atteints de démence. Étant donné la nature relationnelle du soi – que le soi est constitué par nos relations avec les autres – il y a des degrés auxquels les individus peuvent être gênés dans la tâche de développer un sens cohérent de soi. En effet, lorsque l'on est socialement ou mentalement en difficulté, les gens peuvent être « maintenus en place » par d'autres pour préserver un sens de soi. Cette tenue peut être favorable, mais une préoccupation centrale dans ma thèse est qu'elle peut aussi être oppressive. Ma thèse se termine en montrant

que si se sentir coincé dans le temps peut causer plusieurs dommages personnels et relationnels, ces dommages peuvent être surmontés par nos liens avec les autres dans notre cercle social grâce à la confiance interpersonnelle.

Acknowledgements

This thesis would not have been possible without the world's best support team.

Natalie Stoljar and Ian Gold were my academic icons before having the opportunity to study at McGill. After having the honor of working with them, I can now say that they are my personal icons. Natalie has been the person I have leant on for emotional support during the darkest period of my PhD journey and is someone I am in constant admiration of as a woman in philosophy. She has gone above and beyond to support me in every way she can during my PhD, from attending my conferences, encouraging me to apply for scholarships, and meeting me for coffee every week. She also is probably one of the best editors in the *world* and has truly helped me to express my ideas clearly. Ian has also been an incredible support to me, shaping me to be an interdisciplinary scholar and meeting me whenever I need help (which, unfortunately for him, is quite frequently). Ian is the first male philosopher to truly build my self-confidence. His belief in my abilities as a writer and thinker has been so transformative for me that I started to believe in myself at the beginning of my PhD. His encouragement to write and not judge myself has been critical to my creativity. Both Natalie and Ian are academics whose intellectual generosity knows no bounds, and I have been very lucky to benefit from this over the years. Perhaps, most importantly, though, they have shown me how to be an academic with integrity. I am indebted to them for that. However, it is really unfortunate if they think that this submission will result in me no longer desiring to chat with them. That won't ever happen. The sentence has just begun.

I have also been lucky enough to receive substantial financial support from McGill. As someone who had to waitress, bar-tend, and do administrative work throughout her

undergraduate degree, the freedom which comes with receiving funding for four years has helped me substantially. I could not be here without it. And I do not take it for granted. I would therefore like to thank the Department of Philosophy at McGill for providing scholarship funding for my PhD. I am also grateful to have received the Graduate Mobility Award to visit the Centre for Philosophy of Memory at the Université Grenoble Alpes, as well as several scholarship awards from the IHSP for training in health policy and practice, the Storrs McCall award, the McCall MacBain award, the Centre de Recherche en Éthique scholarship and Groupe de Recherche Interuniversitaire sur la Normativité scholarship.

I am indebted to Lisa Bornstein, Laurence Kirmayer, Daniel Weinstock, and Eric Latimer for offering me research assistantships during the course of my PhD to help with costs of living. I have been fortunate to have been taught by and interacted with many fantastic scholars at McGill over the years, to whom I also owe a special debt: Alia Al-Saji, Dirk Schlimm, Elizaveta Solomonova, Michael Hallett, Gaëlle Fiasse, Kristin Voigt, Hasana Sharp, Emily Carson, Marguerite Deslauriers, Stephanie Leary, Jai Shah, Jennifer Proudfoot, all members of The IHSP and IHSP staff, all members of the CRÉ and GRIN research groups. I am grateful to the administrative staff who have helped me out over the years too: Andrew Stoten, Angela Fotopoulos, and Tianyi Zheng.

I dedicated this thesis to two brilliant friends: Sophie Osiecki and the daughter of my brilliant friend Matilde Aliffi. This thesis would not have been possible without drinking sangria with Sophie every other week, and without the notes and gifts Matilde gave me before my attendance at McGill. Both, in different ways, had to leave philosophy. But their impact on me has remained.

Other friends deserve mention in this thesis too. Here at McGill, Yasmin Haddad, Clara Bolster, Matthew Palynchuck, Katie Saulnier, Martina Orlandi, Jordan Walters, Sebastian Rodriguez Duque, Khaleel Rajwani, Davis Collins, Thomas Minguy, Ray Aldred, Valérie Lynn Therrien, Jean-François Rioux, Moritz Bodner, and Walter Reid. At home, my philosophy big sister, Helen Ryland, for her wit, encouragement, teaching me that you can disappear from conferences without an explanation to visit a museum of your choice, and, most importantly, her sarcastic humour, which I share. My best friends Shanice Ennis-Melhado, Erinne Colhoun, and Liam Faller for the constant support, the love, and for reminding me who I am when I forget. My Birmingham alliances, cheerleaders, all-around-legends, Lauren Blackwood, Jodie Lauren Neville, Shannon Oates, Amy Conkerton-Darby, Liberty Severs, Yingna Li, Serena West, Alex Miller Tate, Helen Waite. Professor wise, Scott Sturgeon (*Texan Gandalf*), Alistair Wilson (*Cookie Monster*), Hanna Pickard, Heather Widdows, Alison Jaggar, Matthew Parrott, and Herjeet Marway deserve mention for their love and support over the years.

I'm not sure this completion would be possible without the support of my feminist philosophy students here at McGill. With love and gratitude to Kiran Basra, Alua Kulenova, Riyana Karim, Alyssa Cohen, Kennedy Roberts, Alice Pessoa de Barros, Fion Zhen, and Sequoia Kim, and all those who have written to me over the years. You are my heart; I cannot wait to see where you all go from here.

Finally, the lovely Walsh family. My mother for teaching me how to stand up for myself and other women. My father for teaching me the discipline of hard work. My sister for teaching me the beauty of art and photography. My aunt for believing in my writing. Our dog, Brontë, for teaching me every day that the little moments of love matter most.

My grandma, Doreen, passed away just before I left for Montréal. Doreen was an excellent gardener. She tended to everything and had this incredible insight to know what each flower needed to flourish. I flourished with her love, and I carry that quality in her, and the importance of that quality, everywhere I go.

With an eternal warmth and appreciation in my heart,

Em

Contribution of Authors, References to Original Publications, & Copyright

All chapters of the thesis are authored solely by myself, but I attained permission to include the papers which are published in the body of this thesis.

The first chapter is the only chapter of the thesis yet to be sent out for publication.

The second chapter, “Memory, Colonialism, & Psychiatry: How Collective Memories Underwrite Madness” is set to be published in the journal *Philosophy, Psychiatry, & Psychology* later this year as an original article, which will be accompanied by two commentaries, one from a clinician and the other from a philosopher.

The third chapter, “Cognitive Transformation, Dementia, and the Moral Weight of Advance Directives” is published in the *American Journal of Bioethics*. It is accompanied by 17 open peer commentaries responding to the article, along with my own commentary on the responses.

The fourth chapter, “Trusting at the End: Mosaic Trust and Dementia” is set to be published in the Moral Psychology of Trust edited collection (edited by D. Collins, M. Alfano, and I. Jovanovic) by Lexington Books later in the year.

Madness, Marginalization, & Mental Time Travel:

The Personal and Relational Harms of Feeling Stuck in Time

Introduction

Human thought can move in time, re-live the past, and imagine itself in the future. This capacity has become known to philosophers and cognitive scientists alike as the capacity to mentally time travel (MTT). This thesis will examine the ways in which an individual's capacity for MTT can be impaired by feeling stuck in time. "Stuck in time" is intended to capture the idea that individuals can end up feeling stuck in either memories of their past, their beliefs, or in a set of circumstances they feel they cannot escape. Before delving into the theoretical motivation and consequences of this project, I will begin by explaining my personal motivation for pursuing it.

My initial fascination for the problem of MTT arose due to experiencing my own ability for MTT unravel as a result of having Post-Traumatic Stress Disorder (PTSD) and through my volunteer work with the Alzheimer's Society in England. I began to notice how individuals can end up feeling stuck in time, albeit to more or less severe degrees, and how this experience results in both personal and relational harms. One incident stands out in my mind as the beginning of wondering what those harms might be.

On the way to meeting some of my good friends in Birmingham for a few drinks, I noticed an elderly lady crying and screaming. It was hard to figure out what she was yelling because her words were muddled, and she was so distressed. She insisted she had lost her son and had been looking for him all day. She had phoned the police, but they wouldn't look for him until a specific time period had elapsed. I believed she could have dementia, given the muddled content of what she was saying. I asked her what date she thought it was, after having learnt through my volunteer work that if you find an individual with dementia, it is essential to meet them in the time they believe themselves to be in, so as to not add to their distress. She responded, "9th July 1978". This was alarming, given that we were then at the beginning of 2018. "I'm sure we'll find him… let's look over there…" I said as I typed in the phone number for the Alzheimer's Society office. Before making the call, I saw an elderly gentleman running toward us. He seemed relieved to have found her. "Our son is missing", she told him hysterically. "I've been looking all day". "Our son went missing in 1978, but our son is no longer missing; we found him, remember? He's at home, worrying about you, *who have been missing all day*." He turned to me, "It's a horrible thing, isn't it, dementia? No one understands her. I don't understand her."

I was struck most by how I *did* understand her and others like her whom I had met through my volunteer work. Oddly, I felt connected to her because she showed me a more painful reality than my own, but one that wasn't wholly alien to me. Although I wasn't stuck in 1978, I was stuck in a traumatic event that had occurred the year before, in 2017. That traumatic event, which is elaborated on in Chapter 1 of the thesis, was finding a young man attempting to commit suicide. I wouldn't return to the location thinking it was happening here and now like the lady with dementia. But I did understand how horrifying it is, on a physical and emotional level, to *think* that you are in a time and place which you no longer inhabit. To feel stuck in your ability to move in time. Whilst this occurred in different ways for us both, it placed us both stuck in time, and this feeling of being stuck in time rendered us both in a state of mental distress.

Individuals with semantic dementia have often been referred to as stuck in time because they cannot re-live the past, and project forwards in the future, as their dementia has affected the capacity to mentally time travel (MTT) (Irish, 2016).[1] The same has been claimed for individuals with severe ventromedial damage (Gerrans & Kennett, 2010; 2017). This thesis argues that being stuck in time is far more ubiquitous than the philosophical literature on memory has acknowledged – several social and psychological conditions can impair the ability to MTT. Further, MTT, which has typically been conceived as a capacity that is *in our head*s, is not individualistic but rather is socially and relationally mediated. Philosophers of memory and philosophers of psychiatry must therefore take more seriously the potential for social and relational threats to MTT, which can result in serious mental distress.

In making this argument, I explicate how memory is connected to concerns regarding marginalization and concerns regarding madness.[2] There are many questions of memory that have generated debate in the philosophical literature: the notion of MTT (Michaelian, 2016; Perrin & Michaelian, 2017), the connection between collective memory and individual memory (Michaelian & Sutton, 2019; Isurin, 2017; Ricœur, 2003), questions concerning personal identity and memory (DeGrazia, 1999; Korsgaard, 1988; Parfit, 1971), questions of agency and memory (Gerrans & Kennett, 2017; 2010; Gerrans, 2018; Kennett, 2009), and questions of how memory is tied to social identity (Campbell, 2014; Fivush & Graci, 2017). Yet, an adequate philosophical exploration on the connections between memory, madness, and marginalization is absent from such discussions.

---

[1] We may have reservations that memory is tied to the imagination in this way. My subsequent literature review will show why we have strong evidence to believe memory is reconstructive in this way.

[2] I adopt Rashed's positive conception of the term madness for the purposes of this thesis. Madness in this view is used to symbolize that adverse social, relational, and political conditions can cause individuals to be deemed *mad,* even when the conditions themselves foreground madness.

A Comprehensive Review of the Literature

In the first section of this introduction, I will begin by explaining what I mean by MTT, madness, and marginalization. The second section provides a brief overview of the literature on the relational self that is the background to my argument that one's ability to MTT is socially and relationally mediated. The third section describes three cycles of interrelationships between MTT, madness, and marginalization, including how each chapter explores these themes.

Mental Time Travel

Remembering is of central importance in our daily lives, but why it is vital and what knowledge remembering gives us have been points of philosophical dispute. Early thinkers, Aristotle in particular, conceived of memory as separate from imagination because there appeared a difference between remembering on the one hand and 'expectation' on the other (Bloch, 2007; Michaelian, 2016).[3] Such theories conform to traditional metaphors of memory, which also view memory as distinct from imagining. Metaphors of memory dominated the discourse on memory in the 20th century. These include the idea that memory is like a storehouse, memory is like a photograph, memory works like a library, and memory works as a computer (Draaisma, 2000; Michaelian, 2016; Michaelian & Sutton, 2019). These metaphors highlight that memory is helpful for storage. That is, they emphasize and conceive memory as "being essentially a matter of storing and retrieving fixed items of information" (Michaelian, 2016, p. 7). These metaphors help explain the plausibility of Aristotle's initial thoughts on memory (Bloch, 2007).

---

[3] Expectation here is meant to capture imagining the future. Aristotle, unwittingly, began the suggestion that remembering MTT and future-orientated instances of MTT are different in kind.

On this view, memory is distinguishable from imagination because memory is reliable, whereas imagination is not.

There is no doubt that there is a relationship between remembering and imagining the future, but MTT theorists are still debating the nature of that relationship (Perrin & Michaelian, 2017). There are generally two approaches in this debate. The first, or *discontinuist* approach, suggests that MTT can come apart from future thinking, as remembering and future thinking are mental occurrences of different kinds (Debus, 2014). This echoes the beliefs of early thinkers such as Aristotle, who believed imagination and memory to be different kinds. The second, or *continuist* approach, suggests that there is no fundamental difference between MTT directed towards the past and MTT directed toward the future because the same faculties are engaged in both (Suddendorf & Corballis, 1997; Perrin & Michaelian 2017). This thesis is grounded by continuist theories of MTT (Michaelian, 2021; Suddendorf, Addis, & Corballis, 2009), which endorse the former sense of MTT where MTT is connected to future-directed imagining. I believe it to be most plausible given the wealth of psychological evidence which ties imagining to remembering (Michaelian, 2016). According to continuists, MTT is both backwards-looking and forward-looking, because the faculties for memory and imagining the future are symmetrical.

MTT theorists who view memory as connected to imagining seek to reject the traditional distinction between remembering and imagining. They agree with traditional theorists that the reliability of memory is essential and that memory can play a critical role in information storage (Michaelian, 2016; 2021). They disagree that imagination is not utilized in remembering. MTT theorists emphasize that episodic memory, the kind of memory which allows us to revisit episodes from our personal past (Tulving, 1983), is heavily reconstructive.

The word “reconstructive” captures the idea that we construct scenes of our past in our minds and that this reconstruction is not only a matter of conjuring up the relevant stored file like memory metaphors would mislead us into believing. It involves our imagination as much as our storage retrieval systems. Take, for instance, the kinds of mental occurrences that have to occur for me to remember my dog walk this morning when my mother asks me how it went. I have to think back to the people I saw, the weather, the interactions I had, and my dog's face. According to MTT theorists, all of this involves *constructing* the past episode inside my mind in a way that combines memory and imagination. Michaelian argues, “The episodic memory system is in reality a general episodic construction system, designed to draw on information originating in past experience to simulate possible episodes” (Michaelian, 2016, p. 103). Michaelian notes that this construction system is wholly compatible with memory being reliable and accurate but also tied to imagining, as on such view remembering is imagining the past reliably and accurately.

Reconstructive accounts of memory argue that memory is tied to imagination, and these accounts underpin this thesis. I assume the plausibility of such views given the wealth of psychological research connecting imagination to memory (Perrin & Michaelian, 2017; Michaelian, 2016; Suddendorf & Corballis, 1997; Suddendorf & Corballis, 2007; Suddendorf, Addis, & Corballis, 2009). My accounts in Chapter 1 and Chapter 2 of the thesis also support these theories, showing how memory is connected to imagination, even if we are not yet sure precisely in what ways.

Madness

Negative connotations of the term *madness* are shifting, because of the realization that madness can be a response to a socially unjust world and using the term 'mad' can be a demand for recognition in such a world (Rashed, 2019). The invention of Mad Pride activism has, in more senses than one, "brought mental health activism squarely within the ambit of moral and political discussions on identity and recognition, and have made it clear that what needs to change is far bigger than the profession of psychiatry, it is society itself" (Rashed, 2019, p. 4). The question thus arises: what is it about our social spheres that result in, and perpetuate, instances of mental distress? There are several plausible responses to such a question, including economic conditions like poverty, social discrimination like racism, homophobia and transphobia, as well as personal genetic vulnerability to developing certain mental illnesses. Foucault claims that madness repeats itself as a social phenomenon and serves a social and political goal: the goal of exclusion.[4] He writes:

> Once leprosy had gone, and the figure of the leper was no more than a distant memory, these structures still remained. The game of exclusion would be played again, often in these same places, in an oddly similar fashion two or three centuries later. The role of the leper was to be played by the poor and by the vagrant, by prisoners and by the 'alienated'.
>
> (Foucault, 1988, p. 6)

---

[4] Foucault's point that we forget how madness has worked is powerful in the context and scope of this thesis. Forgetting, after all, is tied to successful remembering: remembering every detail would be neither healthy nor desired (Small, 2021). Yet forgetting is also harmful to the exact point Foucault makes: historical patterns repeat if we fail to recall them and fail to challenge them, and, like Fanon's quote emphasized, these patterns have affective weight in the minds and bodies of marginalized individuals. One of the primary contributions of this thesis is highlighting the connections between memory, madness, and marginalization through the faculty of MTT.

Madness so constituted appears to have three core elements: (i) the category of madness can be deployed to control populations which deviate from social norms (ii) madness can arise from one's social, environmental, and economic conditions if these conditions are distressing; and (iii) madness is relational, insofar as others can make us feel that we are going mad even when we are not. Each element will be a vein seen throughout this thesis, so let us explore them in turn.

An illustration of (i) is the case of female sexuality. If female sexuality was expressed before the 20th century, it was not simply problematic; it was pathologically disordered. This move from problematic to disordered is an essential social move for the reasons Foucault elaborates. Problems can, after all, be solved, whereas disorders must be *treated*. Women were thus thrust into an odd position where they must only sexually desire their husbands but only desire sex when the husband wanted to have sex with them. They were deemed mentally and physically normal if they had little sexual desire (Maines, 1999). Any expressions of disappointment or low interest in sexual activities with husbands who actively wanted to have sex with the woman in question were deemed pathological (Maines, 1999, p. 39). Similarly, various pathological categories of female sexuality were generated, such as hysteria and neurasthenia. As Maines notes, "If the normal functioning of female sexuality was defined as a disease, women must have seemed frail indeed" (Maines, 1999, p. 37). Pathological disorders were thereby invented to try to control and limit women's sexual expression, according to Maines (1999). The result is that deviating from the social norm came with a high price, the price of being mad, and the cost was being socially marginalized and subsequently excluded.

Society influences madness by the adverse social, political, environmental, and economic hardships it can place on individuals' lives (point ii). Consider the rising level of homelessness in the U.S. (Shinn & Khadduri, 2020). Homelessness does not occur at random: homelessness is primarily an urban phenomenon which arises not as a result of the individual deficit but as a result of constant increases in housing prices. Homelessness is also a racialized phenomenon. In Canada, this falls on Indigenous people, who make up 24-34% of the shelter population, and in the U.S., this falls on the African American population, who are at three times the risk (Gaetz et al., 2016; Shinn & Khadduri, 2020). Homeless individuals are at higher risk of being diagnosed as mad because they are in situations which cause them to be "mad," such as trying to survive without money and shelter. Moreover, their economic condition contributes to the madness they experience, either by initiating it or by exacerbating pre-existing mental illness.

The result of this is further social marginalization. Those deemed mad are classed as deviant and often subject to social judgement. This occurs from the widespread following belief:

> The mentally ill are not us – are *other,* different from you and me. But any psychiatrist will tell you that isn't true. Mental illness is just a frayed, weakened version of mental health.
>
> (Gold & Gold, 2015, p. 13)

The belief, however, that the mentally ill are different in a negative way often results in social marginalization. As we have seen, madness is often shaped by social, political, economic conditions, and gendered dimensions. This means that certain groups may be at more risk of being diagnosed with a disorder in the first place, which, whilst known, is not *widely* socially acknowledged.

Borderline personality disorder (BPD) is one such mental illness. Nancy Potter notes that BPD is a gendered phenomenon and that we need to take seriously how women resist their psychiatrists as this could lead to their flourishing in the future (Potter, 2016). Potter's work on the virtue of psychiatric defiance suggests that patients often rebel against their psychiatrists in ways that could eventually lead to their overall flourishing in the unjust societies in which we find ourselves because one often needs to be defiant to survive in unfair conditions. The following conversation between a psychiatrist and their patient is one such instance of the kind of defiance Potter believes could be helpful in the long run.

> Psychiatrist: First of all, you know the rules here. We can't work on your intense feelings when you physically act them out. We need to use words.
>
> Patient: Okay then. Fuck you.
>
> (Reiland, 2004; Potter, 2016)

She notes that flaunting the conventions and hierarchies of the therapeutic relationship may result in flourishing given the social background in which the discussion is taking place. Mental illnesses are thus relational (point iii). They are affected by the relations in which we stand with others, both in the process of becoming unwell and in the process of becoming better.

Marginalization

There are several areas of philosophical work in which marginalization and imagination are connected that have substantially illuminated the arguments of the thesis: feminist philosophy, critical race theory, and work in bioethics on ageing populations. Feminist philosophers have been concerned with how people who are marginalized have their imaginal repertoires constrained (Mackenzie, 2000, p. 124). Imagination is also key to philosophical critique, when feminist philosophers "take a theoretical model that has become so familiar as to be invisible and confront it with undertheorized areas of experience" (Lloyd, 2000, p. 113). Imagination plays a dual role in marginalization, with a negative and a positive side. Feminists caution that we often become what society demands of us, and we acquiesce to these demands without feeling oppressed by them, even though the ideals themselves are oppressive (Widdows, 2018; Garcia, 2018). Yet, feminists can, according to Mackenzie (2000), imagine otherwise for themselves, rejecting such ideals and imagining new possibilities for future action.

Decolonial theorists also emphasize both aspects of imagination. For instance, Frantz Fanon envisioned radically new possibilities and saw the connections between madness, marginalization, and memory, even though he never applied these to MTT, as memory was still not believed to be reconstructive during his time.[5] Fanon's focus on the temporal oppression faced by racialized individuals will be a theme throughout the thesis.

> Surtout n'exigez aucune précision : vous n'en obtiendriez pas. Par exemple, dans les douleurs d'allure ulcéreuse, il est important de connaître l'horaire des algies. Cette conformité aux catégories du temps, le Nord-Africain semble y être hostile. Ce n'est pas incompréhension, car souvent il vient accompagné d'un interprète. On dirait qu'il lui

---

[5] This is why I express in Chapter 2 that Fanon pre-empted new continuist theories of MTT, which suggest that remembering and imagining are interconnected in the same mechanism.

> coûte de retourner là où il n’est pas. Le passé, pour lui, est un passé cuisant. Ce qu’il espère, c’est de ne jamais plus souffrir, de ne jamais plus être face à face avec ce passé. Cette douleur présente, qui mobilise ainsi les muscles de son visage, lui suffit. Il ne comprend pas qu’on veuille lui imposer, même par le souvenir, celle qui n’est déjà plus. Il ne comprend pas pourquoi le médecin lui pose tant de questions.
>
> (Fanon, 1975, p. 12)[6]

Fanon’s influential work, *Le Syndrome Nord Africain,* reveals the psychological and physical costs for racialized individuals who feel incessantly pulled between two times: a past they did not create and a present that causes them to feel mental distress.[7]

Alia Al-Saji’s work, combining feminist theory and Fanon’s work, solidifies the theme of the temporal aspect of oppression. As Al-Saji notes, “[r]acism and colonialism are temporal formations, as well as being geographical, economic, social, and imaginary ones. They manage not only territories and bodies, but also histories, pasts, and futures” (Al-Saji, 2021, p. 177). This can be seen most explicitly in Chapter 2 of the thesis, where I explore the impact of racialization on both episodic memory, those that individuals have regarding their own pasts, and collective memories, the shared narratives we hold in our social milieu, particularly in relation to the other main themes of madness and memory.

---

[6] English translation*:* Above all, do not demand any precision: you will not obtain any. For example, in the pains of an ulcer, it is important to know the hour the illness began. It is this conformity to the categories of time that the North African seems to be hostile. It is not incomprehension, because often he comes accompanied by an interpreter. One used to say that it cost him to return there where he is not anymore. The past, for him, is a burning past. What he hopes is that he never suffers anymore, that he never has to be face to face with this past. This present pain, which also mobilizes the muscles in his face, is enough. He does not understand that one would want to impose on him, even by way of memory, that which has already passed. He does not understand why the doctor poses so many questions to him (Fanon 1975, 12).

[7] Why this is will be explored more carefully in Chapter 2 of this thesis.

Marginalization is also experienced heavily by ageing populations, who are often subject to negative stereotyping regarding them as “necessarily dependent, frail, out of touch, or a burden” (Jecker, 2020, p. 240). This can result in persecution, infantilization, and, at the extreme, death resulting from marginalization. Consider responses from people regarding whether they would hire an individual of a certain age for an occupational position and such people responding that they would not (Jecker, 2020) or simply by witnessing the horrifying aspects of how ageing populations were treated during the Coronavirus pandemic in many parts of the world. Just as gender, sexuality, and race affect one’s self-perception through internalizing social stereotypes, so too does ageing, which risks increasing one’s chances of early mortality (Levy et al., 2009). In addition, elderly persons facing cognitive deficits suffer a more insidious kind of marginalization, because of judgements by others that such lives may not be worth living (Brison, 2021).

The Relational Self

I now turn to an important background theme of the thesis, that of the relational self. Recall that traditional pictures of MTT have tended to be individualistic, insofar as they describe one’s ability to MTT as being in one’s own head. Feminist philosophers and others call the individualistic notion the “atomistic” self. Philosophers who defend the relational self, in contrast, argue in favor of the view that selves exist *only* in relation to other selves (Brison, 2017). According to relational accounts of the self, personal identity is “constituted in and through one’s lived bodily engagement with the world and with others” (Mackenzie, 2009, p.119). But why is the relational self a more plausible, realistic view of the self than traditional individualistic accounts? Why ought philosophers of mind endorse it themselves?

The first reason is transparent: humans are social creatures with social capacities. There is at the very least a sense in which we would not have developed as the persons we are in the absence of relations to others. As Annette Baier puts it, we are all “second persons” (Baier, 1985). Second, if one’s social capacities are not utilized, and one’s social needs are not met, one suffers mentally and physically (Motta, 2021), suggesting that retaining a coherent sense of self requires being in contact with others. Relationality fulfils our inherent social needs of attachment and social integration. Third, we are social beings with moral obligations to others in our social spheres. The obsession with the individual obscures these social, ethical obligations and ignores that “we are responsible for the meaning of each other’s inner lives” (Scheman, 1983, p. 241). Fourth, a social-relational, narrative view of the self is widely endorsed in both psychiatric contexts (Groleau, Young, & Kirmayer, 2006; Kirmayer, 1996) and philosophical discussions of the self (Atkins & Mackenzie, 2008).

To make sense of our experiences in time, we cannot treat them as isolated episodes but rather need to place them within a narrative structure that itself relies on social relations for its production and reproduction. Brison summarizes: “How *other* people use words constrains our self-narratives” (Brison, 2017, p. 227). For instance, the narrative self is often constrained and manipulated by “master narratives.” As outlined by Lindemann (Lindemann, 2009; 2014), these are familiar stories which permeate the fabric of our self-narrative. Some master narratives, such as stories of family history, are not oppressive in themselves. However, Lindemann argues that certain master narratives are necessarily oppressive, such as “hateful or dismissive master narratives used by the members of a dominant social group to justify their oppression of another, less powerful group” (Lindemann, 2009, p. 420). The effect of the master narrative is particularly explicit in Du Bois’s notion of double consciousness: “It is a peculiar sensation, this double-consciousness, this sense of always looking at one's self through the eyes of others, of measuring one's soul by the tape of a world that looks on in amused contempt and pity" (Du Bois, 2018, p. 8). This "doubleness" results in the sense of both affective and personal disillusionment, where one's sense of self is disrupted and one's self-narrative cannot be reconciled with the oppressive norms imposed from outside.

The concept of the relational self helps to explain why positive social relations bolster one's sense of self and self-esteem, whereas, in negative ones, the sense of self is worn away. In particular, feminist philosophers have focused on how self-regarding attitudes such as self-trust and self-respect can be undermined in oppressive social settings.

Identifying the Three Cycles: MTT, Madness, and Marginalization

Whilst I propose the oversight of the connection between MTT, madness, and marginalization is a mistake, it is an oversight which makes sense. [8] At first glance, after all, it is not easy to see how these concepts are linked. What exactly does marginalization have to do with memory and madness and vice versa? My answer to this question lies in MTT. I argue that the engagement of memory and imagination in MTT can be impacted by madness and marginalization and that there are three interconnected cycles, as follows

(i) madness → impairments to MTT → marginalization
(ii) marginalization → impairments to MTT → madness

---

[8] It is important to note that feminists have not ignored the connection between marginalization and imagination in the same ways that philosophers of psychiatry and philosophers of memory appear to (Mackenzie, 2000; Lloyd, 2000). They simply have not made the explicit connections between MTT, madness, and marginalization which will be the focus of this thesis.

(iii) impairments to MTT → madness → marginalization

The first cycle begins with an individual living with madness, suffering from impairments to their ability to MTT as a result, and subsequently being socially marginalized. The second cycle occurs when negative collective memories, the intersubjectively shared narratives we possess in our societies, once internalized, cause individuals to suffer from impairments to MTT. In so doing, I suggest such individuals are socially marginalized and thus at higher risk of being deemed mad. The final cycle commences once one faces impairments to MTT in the form of memory loss, which results in the diagnosis of a disorder, which in turn can result in social marginalization. The novel contribution of this thesis is to suggest that all these cycles, though different at face value, have a common thread. That thread is that each of these cycles results in the experience of feeling stuck in time.[9] My thesis examines how psychological and social conditions lead to the experience of being stuck and thereby threaten one's ability to MTT.

Chapter 1 begins with cycle (i). It provides an account of the phenomenology of people deemed "mad" or dysfunctional in some way such as those suffering from PTSD.[10] I identify specific dimensions of the capacity for MTT by elaborating how feeling stuck in time results in *impairments* to MTT: the specific dimensions are those of attention, narrative, semantic memory, imaginal possibilities, cognitive flexibility, and self-trust. I argue that accounts of MTT in the current literature are overly individualistic and overlook these features; they fail to consider how components of MTT are relationally and socially mediated. By examining the case of trauma, I argue that we get a clearer relational picture of both what MTT is and why it

---

[9] This account of being stuck in time suggests that individuals can be more or less stuck, and more or less impaired with respect to their ability to MTT as a result. It thus calls into question views which paint MTT as an all-or-nothing capacity.

[10] This is not to generalize that everyone has the same phenomenology when they feel psychologically stuck; this is just to present *one* account of this psychological phenomenon.

matters. The result is intended to capture the cycle that suffering from mental illness often involves impairments to MTT which can then lead to social marginalization.

While Chapter 1 shows that the phenomenology of feeling stuck in time can result from psychological distress (such as PTSD), Chapter 2 argues that it can also result from occupying a marginalized social position. This covers cycle (ii). Utilizing the work of Frantz Fanon, I show how the process of racialization, the process by which hierarchies and differences are made between races (Al-Saji, 2013), is enacted through collective memories. My account shows how negative collective memories underpinned by colonial social structures can factor into MTT for racialized individuals, acting to cut off their imaginal possibilities. Given the interaction between dominant collective memories and MTT, I argue that this places racialized individuals at an increased risk of being diagnosed as mad in our societies.

After examining the psychological sense of feeling stuck in time and the social sense of feeling stuck in time, Chapter 3 and Chapter 4 discusses a marginal group which occupies both categories simultaneously: individuals living with dementia. So doing covers cycle (iii). Chapter 3 explores one social result of living with dementia, namely the individual’s current voice is ignored. As a result, such individuals can suffer from marginalization due to their failing memory. They can be negatively held in their old identity by family and friends who refuse to acknowledge the individual's transformation. I suggest that this is particularly the case when the individual living with dementia is starkly different to the individual the family and friends once knew.

Chapter 4 explores a different angle of the same cycle: an example of what I term “positive holding.” Consider the depiction of Iris Murdoch’s dementia in the film *Iris* (Eyre, 2001). Iris is undoubtedly stuck in time due to her dementia, yet her husband John’s love for her helps to remedy this state in several ways. This chapter suggests that our relations with others, particularly our trust in others, can be one means of transcending the state of feeling stuck in time, even if it is due to a neurodegenerative condition like dementia.

I conclude by offering insights into what the relational and personal harms of feeling “stuck in time” are. I draw out three personal harms from the chapters, which include, but are not limited to: disruptions to self-trust, having a sense of myopic imaginal possibilities, feeling unsafe in one’s body. I draw out two relational harms of these experiences: resistance may cause a cycle of further marginalization, and blunted affectivity can affect one’s relationships.

Chapter 1: The Phenomenology of Feeling Stuck in Time – MTT & Trauma

Chapter one begins our investigation into the personal and relational harms of feeling stuck in time by exploring a psychological threat to mental time travel (MTT) posed specifically by instances of trauma. MTT, the concept that individuals can re-write their pasts and project themselves forwards in the future, has gained significant attention from cognitive neuroscientists and philosophers alike. This attention has resulted in what I call an *individualistic and rationalistic* picture of MTT, which views MTT as an individualistic mechanism that one has the capacity for or lacks the capacity for. On such accounts, MTT is valued for reasons of planning and reasoning, as MTT is useful for both purposes.

I suggest in this chapter that exploring more ubiquitous cases of psychological threats posed to MTT, as in the case of trauma, is theoretically valuable because it challenges the individualistic and rationalistic picture of MTT. Trauma shows us that one's capacity for MTT is not all-or-nothing, one's system can malfunction in degrees, and therefore one's capacity for MTT can exist in degrees. Trauma cases highlight that there are social and psychological elements to MTT, as evidenced by the fact that social and psychological threats can cause impairments to MTT. I show how trauma makes one feel *stuck in time* as one feels trapped in a traumatic memory repertoire. By offering the precise ways trauma disrupts MTT, I aim to give a fuller account of the cognitive, affective, and relational elements involved in what is required for MTT to be successful. This chapter identifies five essential additional elements to being able to successfully MTT: attention, narrative, semantic memory, cognitive flexibility, and self-trust. I suggest that examining these dimensions show us that the current individualistic and rationalistic picture presented in the literature on MTT is incomplete because it ignores these emotional, affective, and cognitive dimensions of MTT.

In drawing out these additional features of MTT that the individualistic accounts fail to mention, I seek to create a new theory of MTT which recognizes that MTT is a relationally and socially mediated capacity; it is not merely an individualistic one, and its only value is not simply rationalistic. The result is that it can be undermined subtly in ubiquitous everyday encounters, by social and relational threats.

Exploring the case of trauma maps onto exploring cycle (i) of madness, MTT impairment, marginalization. It illustrates that whilst we may conceive of psychological threats as being inner to the individual, they are not, because threats of trauma usually occur from social and relational sources.

§1 Introduction

The capacity to "rewrite" one's past and imagine one's future, known as the capacity for mental time travel (MTT), has received growing attention from philosophers and cognitive scientists alike (Fivush & Graci, 2017; Gerrans & Kennett, 2010, 2017; Gerrans & Sander, 2014; Michaelian, 2016; Michaelian & Sutton, 2019; Perrin & Michaelian, 2017; Suddendorf et al., 2009; Suddendorf & Corballis, 1997). An influential account in the field of MTT is simulation theory, where remembering is a matter of imagining the past reliably and with a degree of accuracy (Michaelian, 2016, 2021). This theory argues that there is a close connection between imagining the future and imagining the past. It is influential and widely endorsed by psychologists and philosophers of memory on the basis of the established connection between imagining and remembering.[11] This theory of MTT will underpin this article. MTT theorists who endorse the simulation view claim that just as we reconstruct the past by re-imagining it and mentally re-living it, we do the same for future possibilities. As Michaelian notes, "when the target of the system is, for example, a future episode, simulation cannot amount to reproducing a previous experience of the episode; instead, simulation is a matter of drawing on a range of past experiences to produce a novel representation of the target episode" (Michaelian, 2016, p.103). The same applies to memories in one's past: when one remembers, one is conjuring up potential versions of what happened by drawing on data from imagination and memory to do so.

MTT, so construed, is essential to effective agency (Gerrans & Kennett, 2010, 2017; Kennett, 2009).[12] This is because agency requires engagement with a future self and the capacity to reason based on what would be best for that future self. That demands employing memories of the past to imagine different scenarios for that future self. For instance, if I am deciding whether or not to buy a house or rent an apartment, I will need to recall memories from my past, such as my financial resources and my preferences, to simulate possibilities of which option would be better for me. Without having evidence from past experience, it is challenging to engage in future planning and, therefore, to make these kinds of daily decisions.

Specific pathologies demonstrate that the faculty of MTT is crucial to daily functioning. Individuals with severe memory impairments reveal the harm caused by an inability to MTT. One example that has prompted subsequent discussions in philosophy and cognitive

---

[11] Despite its influence, this theory is disputed by discontinuists who suggest a difference between future-directed MTT and MTT used for remembering. For a rich discussion on the differences between continuist and discontinuist accounts of MTT, see (Perrin & Michaelian, 2017)

[12] Gerrans and Kennett focus on moral agency in particular. The discussion of moral responsibility and mental illness is beyond the scope of the article, so I focus exclusively on their claims regarding agency more broadly defined.

neuroscience is that of Clive Wearing, a British musician. He was infected with the herpes simplex virus at 46, which demolished his hippocampus (Suddendorf et al., 2009). Wearing retained his vocabulary and his abilities to do things such as play the piano, conduct a choir and recall the faces of his wife and family. However, he lost his episodic memory altogether, forgetting what he discussed a few moments ago. Wearing's amnesia left him in a state where "[h]e is continually under the impression that he has just woken up, or recovered from being dead. His conscious experience is entirely of the present" (Suddendorf et al., 2009, p.1). Wearing could not relive the past and imagine the future, resulting in agentic losses. Philosophers have explored some of the agentic harms that arise when individuals with severe memory impairment feel stuck in an eternal present, notably that such individuals are severely compromised in their ability to plan and reason about their future (Gerrans & Kennett, 2010, 2017).

The discussion about MTT and mental illness so far has focused on either severe memory impairment, such as cases like Wearing, where one's ability to MTT is completely eroded (Gerrans & Kennett, 2010) or on confabulation, where one's episodic construction system used in MTT is taken to be malfunctioning so that one's ability to imagine the past reliably and accurately is severely compromised (Michaelian, 2020, 2021). Both kinds of discussion presuppose a dichotomy between those with full capacity to use MTT and those without it. Drawing such dichotomies, whether explicitly or implicitly, is problematic because this ignores how MTT can be impaired in degrees. It paints MTT in an all-or-nothing, black and white light.

I contend that these accounts are incomplete because they share the same individualistic, rationalistic kernel at the center of their accounts, which views one's capacity for MTT as a mechanism one either has in full or lacks. The kernel is individualistic insofar as it views MTT as internal, it is in one’s head, and one’s capacity for it is affected by neurological or psychological deficits. It is rationalistic because it emphasizes MTT as valuable for reasoning, planning, and action. I propose that this lens gives us a limited picture of what MTT is and its purpose. To provide a fuller picture of MTT, I suggest a relational picture of MTT needs to be painted. The reason for this move is two-fold. MTT is an internal capacity insofar as it is in the head; yet it is still a socially and relationally mediated capacity because many brain mechanisms are (Kirmayer et al., 2022). Therefore, offering a view of MTT as wholly inner is, at best, incomplete. Similarly, the insistence that MTT’s value lies in planning and reasoning obscures that individuals may have impairments to MTT, which do not impact this but still result in mental distress. Examining the case of trauma problematizes the individualistic and rationalistic picture (from herein, IRP) of MTT, which has conceptualized it as an all-or-nothing capacity essential for primarily planning and reasoning. Significantly, cases of trauma show *how* one’s system can be threatened and can malfunction. These cases move us away from the individualistic picture of MTT by showing that the psychological cannot be separated from the relational and the social, as relational and social threats can undermine one’s capacity for MTT.

This article has three related aims. The first is to broaden the range of philosophically interesting cases of MTT impairment, beyond looking at issues of severe memory impairment and confabulation, and to elaborate a fuller account of MTT. I explore a case where MTT malfunctions in degrees but has not been erased, where one’s sense of reliability and accuracy is beginning to fade, but one is not stuck in one’s imagination as in cases of confabulation. The second aim is to argue that the IRP of MTT, whilst conceptually and theoretically valuable, is in its present form incomplete. It is incomplete because there are significant cognitive, affective, and emotional elements to MTT which are overlooked in the IRP. Finally, the article argues that how MTT is relationally and socially mediated demands greater philosophical attention. I build on research that explores the affective element of MTT and its relation to mental illness (Gerrans, 2018) as well as research on how narrative identity is a socioemotional construction (Fivush & Graci, 2017). The article also develops a new suggestion: MTT can be impaired by a wide range of social and psychological conditions that are ubiquitous in daily life, which do undermine one's capacity for MTT, but do not erode it entirely. It also emphasizes that to understand MTT fully, a dialogue between philosophers of memory, philosophers of psychiatry, and feminist theorists is necessary. These areas are not separate: each is integral to the other, and we cannot address degradations in the capacity for MTT if there is little communication between them.

The structure of this article will be as follows. In §2, I articulate what I call the IRP picture of MTT and its role in agency. This picture emphasizes an all-or-nothing capacity of the individual rather than one that is relationally situated and which comes in degrees. I propose that we need to revise this picture to give a richer conceptualization of the workings of MTT. To do so, I suggest that we observe how MTT malfunctions rather than how it can be eroded entirely, as has been the current focus in the literature (Gerrans & Kennett, 2010, 2017; Suddendorf et al., 2009).[13] My primary interest and focus are on the phenomenology of feeling "stuck in time": where individuals feel stuck in a limited traumatic memory repertoire, a set of beliefs or a set of circumstances. Being stuck in time is a clear malfunctioning of the capacity for MTT which is intended to help one move in time, back to the past, and forwards to the future. I examine a psychological case of feeling stuck in time by elaborating on instances of trauma. Trauma problematizes MTT as an all-or-nothing capacity implied by the IRP of MTT.

---

[13] With the notable exception of conversations on confabulation, which are beginning to speculate how one's ability to MTT malfunctions, albeit in a black-and-white manner (Michaelian, 2020).

It also suggests that more ingredients are essential to successful MTT than the IRP would have us believe. §3 will explain the symptomology of trauma and the range of traumatic experiences one may experience in a lifetime. §4 will unpack the phenomenology of trauma, highlighting the ways trauma impacts one's ability to MTT by making one feel psychologically stuck. I argue that trauma disrupts the following ingredients of MTT: attention, narrative, semantic memory, imaginal possibilities, cognitive flexibility, and disruptions to self-trust. This shows us that, at best, the IRP of MTT is incomplete as it stands. §5 expands on what the revised theory of MTT looks like, what it emphasizes as necessary, and how it affects our understanding of MTT. I conclude by suggesting that MTT is far more relationally and socially mediated than current accounts consider. Thus, it is far more open to social and relational threats.

§2 The Individualistic and Rationalistic Picture (IRP): What MTT is and Why it Matters to Agency

This section will examine accounts of MTT in the current philosophical and cognitive science literature. I argue that these accounts paint MTT in an individualistic and rationalistic light. They consider MTT to be an individual capacity that is fundamental to agency because it allows rational planning. My view problematizes the implications of these accounts, in particular that the capacity for MTT is individualistic simply because it is in one's head, that MTT is all-or-nothing and that MTT is vital for agency for the primary reason that it is essential to planning and reasoning. I challenge the emphasis of MTT as individualistic by showing how it is relationally and socially mediated. I challenge the claim that it is all-or-nothing by showing how MTT can be impaired in degrees and exist in degrees. Finally, I challenge the view that MTT is vital to agency for reasons of rational planning alone by showing that traumatized individuals may retain this ability but nonetheless can still be impaired in their capacity for MTT and therefore in their agency.

Suddendorf and Corballis argue that MTT is a uniquely human capacity, which begins to be evident in neurotypical subjects at around four years of age. They suggest that, unlike humans, species such as great apes are "confined to a 'present' that is limited by their current drive states" (Suddendorf & Corballis, 1997, p.1). Humans, however, can transcend their current drive states by employing MTT. In their view, MTT is a flexible process of reconstruction which necessitates the use of specific cognitive capacities. Suddendorf and Corballis suggest several preconditions that allow humans to exercise MTT. These ingredients are a level of self-awareness and self-knowledge, an imagination capable of temporal reconstruction, an ability to gain knowledge by experiencing the world around oneself and have a minimal ability to distinguish false beliefs from true beliefs, an ability to meta-represent one's mental states, to attribute past mental states to one's earlier self, and an ability to disassociate from current mental states, insofar as one needs to be able to transcend the present moment to MTT on their view. They point out that animals lack some of these abilities, and so do children with autism and children below three and a half years of age. All fail to be able to disconnect from the present moment, which is a prerequisite for MTT in their view. Suddendorf and Corballis hypothesize that being able to MTT into the future appears to rely on the exact mechanism that being able to MTT into the past does, as well as similar preconditions.

Suddendorf and Corballis suggest that "[t]ime travel into the future is in a sense an extrapolation from time travel into the past, similarly involving the ability to escape the influence of the current mental state" (Suddendorf and Corballis, 1997, p. 8). Michaelian, synthesising evidence from psychology, goes one step further. He defends a theory of simulation which suggests that imagining and memory cannot be separated because to remember is to imagine the past reliably and accurately (Michaelian, 2016, 2020, 2021). Michaelian's theory argues that S remembers an event *e* if:

- S now has a representation R of *e*
- R is produced by a properly functioning episodic construction system which aims to produce a representation of an episode belonging to S's personal past.

(Michaelian, 2016, p.107)

This insistence on the correct functioning of the episodic construction system is necessary to distinguish accurate remembering via MTT from imaginal remembering and, further, from defective imagining, as in the case of confabulation. A properly functioning episodic construction system needs to distinguish counterfactual scenarios from factual scenarios and past from future scenarios. It must aim to produce a more-or-less accurate representation of S's personal past, and this system must be reliable in producing these representations. According to Michaelian, the system malfunctions when it fails to achieve this.[14]

Combining insights from Suddendorf and Corballis, along with Michaelian, an individualistic picture of MTT is painted. I call it *individualistic* as it only considers MTT as a capacity in the brain, neglecting to account for the ways in which the brain is impacted by its social situation and failing to see how one's social situation could make one's episodic construction system weaker or stronger. This occurs, I suggest, by their characterization of MTT as wholly 'inner'

---

[14] Michaelian's account of what constitutes a properly functioning episodic construction system is conceptually underdeveloped alone, but, once combined with Suddendorf and Corballis's conditions, this becomes much richer in my view.

to an individual. This combined view emphasizes that one's episodic construction system must be properly functioning insofar as it must generally produce accurate and reliable representations of the past.[15] One's system is likely to produce accurate and reliable representations of the past *iff* they meet the preconditions for successful MTT specified above.

Yet, for a complete account of MTT, we also need to explain why MTT matters for agency. Gerrans and Kennett fulfil this task (Gerrans & Kennett, 2010; 2017), arguing that what is at stake when one's ability to MTT is lost or compromised are several elements critical to agency. Gerrans and Kennett suggest that if moral judgement requires moral agency, as is standardly assumed in the philosophical literature, then deficits in a capacity for moral judgement ought to be reconceptualized as impairments of agency.[16] They argue that deficits in one's ability to engage in MTT, in particular, ought to be reconceptualized in this way. They argue that "A moral agent needs to be able to conceive of herself as a temporally extended entity as a necessary condition for moral reflection and decision-making" (Gerrans & Kennett, 2010, p. 588). Their central focus is to show that MTT is necessary for allowing an agent to respond to moral demands. They argue that MTT affects two critical abilities central to one's agency: planning and rational deliberation.

To illustrate, the authors expand on the Iowa Gambling Task (Damasio et al., 1991) which was created to probe the deficits in decision-making suffered by brain-damaged patients with a lesion of the ventromedial prefrontal cortex. In the gambling task, subjects, starting with a sum

---

[15] It is important to note that Michaelian (2021) is not committed to the view that one must always represent the past accurately to have a properly functioning system. He points out that even individuals who have such a system can regularly misremember details of the past. What is more critical in his view is that individuals reliably represent the past accurately more often than not.

[16] Whilst their argument is concerned with moral responsibility and judgement, I will be focusing on their claims regarding the relation between MTT and agency. However, I hope this discussion will stimulate discussions on the relationship between MTT, agency, and moral responsibility in the context of mental illness.

of money drawn by the experimenter, draw 100 times from four decks of cards which prescribe a monetary outcome. The experimenters discovered that neurotypical subjects tend to develop a preference for the rewarding schedule *before* articulating why they prefer it. However, the same preference is not made by those who suffer from brain damage to the ventromedial prefrontal cortex, who have impairments in their ability to engage in MTT. These subjects can learn the schedules but remain in the "paradoxical situation of being able to understand intellectually that some decks are disadvantageous but nonetheless choosing disadvantageous decks" (Gerrans & Kennett, 2010, pp. 591-592).

Patients with ventromedial prefrontal cortex damage thus motivate the argument put forward by Gerrans and Kennett because these patients quite clearly struggle to plan and reason based on what they know. Their argument is compelling given the evidence that MTT is intimately connected to both planning and rational deliberation capacities. However, by focusing on extreme cases of memory impairment and implying that MTT is valuable to agency only insofar as it is needed for planning and reasoning, they miss the subtle ways in which other impairments to MTT, which are less severe, also threaten agency.[17]

To summarize, the combination of these views paints an IRP of MTT. It is individualistic because it emphasizes MTT as a capacity in one's head, failing to account for the ways one's social situation could affect one's capacity for MTT, and rationalistic because it suggests the

---

[17] The authors quickly dismiss the possibility that mental illness involves impairments in MTT. They articulate, "Of course there are a number of different ways in which an adult person's agency may be impaired, not all of which turn upon failures of MTT. An individual is a grossly impaired agent if the plans and decisions she arrives at in deliberation are constantly overwhelmed by compulsive desires or if she is so depressed that she cannot bring herself to act at all. In these kinds of cases there may be no failure of the capacity for mental time travel" (Gerrans & Kennett, 2010, p.603). I suggest we pause here to explore the possibility that mental distress *does or, at the very least, could* impact one's ability to MTT. The reasons why will become clearer in §4.

value of MTT lies in its use for planning and rational deliberation. I argue that this picture of MTT is incomplete in several ways. It is incomplete because the essential conditions it does specify do not take into consideration the ways in which these conditions are relationally and socially mediated. It problematically implies, in its selection of cases, a dichotomy between individuals who have the capacity for MTT and those who have malfunctioning systems. This is a mistake, I will argue, because individuals can be more or less impaired in several key areas. In addition, its exclusive focus on planning and rational deliberation obscures impairments to MTT that do not threaten planning and reasoning capacities.

§3 Trauma

Before explaining why trauma is a paradigm case of feeling stuck in time, and why this phenomenology can help to better understand additional essential elements to MTT that the IRP overlooks, it is necessary to unpack the lived experience of trauma. Traditional conceptions of trauma posit that trauma is a condition that follows as a consequence of a traumatic event. Traumatic events correspond to a wide range of phenomena including, but not limited to: ongoing conflict in one's place of residence, exposure to murder and death, intimate partner violence, sexual assault, violations of basic human rights, still-born births, excessive financial stress and insecurity, witnessing war crimes and acts of torture, being involved in accidents, and the aftermath of experiencing climate disasters. This is not to say that all those who witness traumatic events have traumatic responses after the fact, as "[o]nly a minority of those exposed to traumatic events (approximately 24%) go on to experience clinically significant symptoms of Post-Traumatic Stress Disorder (PTSD)" (Breslau, et al., 1991; Gilmoor et al., 2020, p.456).

The *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) diagnosis of PTSD has eight elements (American Psychiatric Association, 2013). The first is a criterion of exposure (I). In order to receive a diagnosis of PTSD, one must have been exposed in one of the following ways to a traumatic event: directly exposed and involved, involved as a witness, exposed through an intimate relationship with the person (such as being a family member of someone who has had a traumatic experience), indirect exposure to aversive details of trauma (such as those experienced by first responders and medics). The second criterion is that of experiencing intrusive symptoms (II). Such symptoms can involve experiences like unwanted nightmares, intrusive memories of the event, emotional distress or physical reactivity. The third criterion is avoidance (III). This criterion reflects a want or need for the person affected by PTSD to avoid potential external stimuli which could result in similar experiences, and an attempt to ignore trauma-related feelings and painful emotions. The fourth criterion is negative alterations in cognitive functioning and mood (IV). Negative alterations in functioning and mood involves symptoms such as overly negative assumptions about the world and oneself, exaggerated self-blame or blame of others, feelings of isolation, and difficulty feeling positive affect. The fifth criterion is alterations in arousal and reactivity (V). Symptoms include increased irritability or aggression, difficulty sleeping, difficulty concentrating, heightened startle reaction, engaging in risky or self-destructive behavior, hypervigilance of world and others. The sixth criterion is a criterion of duration (VI): the individual must have experienced ongoing symptoms for a month to receive a diagnosis of PTSD. The seventh criterion is that of functional impairment (VII). In order to receive a diagnosis, at least one of each of the symptoms listed in criterion I-VI and at least two symptoms in categories IV-V must be experienced by the individual. An exclusion criterion (VIII) must also be met: one's symptoms cannot arise as a result of medication, substance abuse, or other illness.[18]

Both the exclusion criterion and the time requirement have been critiqued. The former has received critique because it makes it difficult for particularly vulnerable populations (such as homeless populations who often suffers from addictive behaviors) to receive a diagnosis of PTSD (Gilmoor et al., 2020). The latter has received criticism because some PTSD is so severe immediately afterwards that many take, or at the very least consider taking, their own lives within one month afterwards, particularly victims of gender violence (Campbell et al., 2009).

Cultural context can affect not only the idioms of mental distress used to capture the experience of PTSD but also the source of one's shame around experiencing PTSD. Whilst the question "why is this happening to me?" is a common thread in trauma narratives, the way it is expressed is culturally variable. For instance, a trauma survivor in India might say "I would think to myself, '[w]hy is this happening to me? Is it because my parents married [outside their caste]?" (Gilmoor et al., 2020, p. 458). Often, social marginalization compounds the aftermath of traumatic events and the likelihood that one will suffer PTSD. Those without adequate shelter, and protection, report feeling at constant threat of not only assault, abuse, discrimination but also violence: "[f]or women [on the streets] it is not safe, anybody will come and go… There is no safe drinking water, no safe home…any man can rape her and do something to her…" (Gilmoor et al., 2020, p. 460).

---

[18] Trauma often involves a person experiencing a range of behaviors that correspond to both hyperactivity and hypo-activity. Those who suffer from PTSD symptomology will often express not wanting to eat, not wanting to socialize, not wanting to engage in activities that used to bring them joy, whilst, at the same time, remaining on constant guard against stimuli that resembles the original trauma stimuli. This dual nature of experiencing PTSD can result in a variety of coping strategies, involving both positive strategies and negative strategies. Positive strategies include: distracting oneself, relying on social support, seeking knowledge and awareness. Negative strategies include: engaging in self-destructive behaviors, consumption of alcohol or drugs to 'numb oneself', engaging in sexual promiscuity, escapism, or disconnecting from others.

Thus, although traumatic experiences have common threads, which will be the focus of this article hereafter, there are different lived affects felt according to the different kinds of trauma that is experienced. The nature of the event itself mediates not only the symptoms but the locus of the distress. My focus is on the common threads which knit these experiences together, but it is important to acknowledge the ways in which the experiences differ in order to understand the complexities of living with the traumatic experience.

§4 The Phenomenology of Feeling Stuck in Time

On the 11th June 2017, in Kings Heath Park, I found a young man of around sixteen perched against a tree, attempting to commit suicide, by putting a black bag over his head attached to a canister of helium. Once I realized what I was witnessing, I tugged at the black bag to get it off and found the young man's mouth foaming with blood. CPR worked to help revive him. Several people helped to monitor his breathing until the ambulance crew came. The police phoned afterwards to say that he would be staying in hospital but that he would survive. That day passed, but the memories of that event stuck.

My lived experience of post-traumatic stress disorder (PTSD) reveals a feature of interest which has not been elaborated in the philosophical literature on agency and mental time travel: I felt psychologically 'stuck' in a limited memory repertoire for the year after this traumatic event.[19] The individualistic accounts of MTT that I described in §2 often conceptualize MTT

---

[19] I have focused in this article on what happens if one is preoccupied mentally with the past to the extent that one feels stuck in it. But my account suggests that being stuck in either the present, the future, or one's imagination will also result in impairments to MTT and will likely result in mental distress which will manifest differently in each case. It would be interesting for philosophers of memory and philosophers of psychiatry to explore this consideration further.

as a capacity one either has or doesn't. Yet, my ability to MTT was damaged in a more limited way by this traumatic event.[20] This section identifies the ways in which trauma caused my capacity for MTT to malfunction. I aim to critique the IRP outlined in §2 by fleshing out additional cognitive, affective, and relational components which are essential to MTT. I describe five ways one can feel stuck in time and how this interferes with one's capacity for MTT. The upshot of this section is to show how one's capacity for MTT is relationally and socially mediated.

Attention

> At the time it was very simple, I put the memory of that morning inside a large jar. I took this jar and carried it down, down, down, flights and flights of stairs, placing it inside a cabinet, locking it away, and walking briskly up the stairs to continue with the life I had built…Every time I thought of that morning, another jar was born. Now jars filled every inch of my mind. I had nowhere to put them. They cluttered the stairwells, could not be contained in cabinets. I was full of these sealed jars, no room to sit or walk or breathe.
>
> (Miller, 2020, pp. 26, 32)

Attention is important to be able to engage in MTT successfully. As explained in §1, MTT demands weighing up and imagining different future options for oneself. If one's attention is severely diverted, this will inevitably affect one's ability to engage in MTT. Miller's metaphor captures an important common feature of traumatic experiences which shows how traumatic

---

[20] This, therefore, differs from the conception of being stuck in time which exists in relation to individuals who lack the capacity for MTT altogether such as those with semantic dementia (Irish, 2016) or ventromedial damage (Gerrans and Kennett, 2010).

experiences can consume one's attention: the attempt to turn attention away from the set of memories involved in the traumatic experience is often futile. The more one tries to ignore the traumatic experience, the more salient the experience becomes and the more it demands affective and cognitive attention. My own experience of PTSD was in line with this: the more I tried to ignore the new memory repertoire the traumatic experience produced, by cognitively and affectively burying it and refusing to speak of it, the more it would take up my mental, physical, and emotional space. Both neuropsychological and psychotherapeutic research demonstrates that persons who experience PTSD focus their attention on trauma-associated stimuli, as the trauma-associated stimuli are cognitively salient and highly valued by the person affected, causing the stimuli to demand greater cognitive attention. This generally has two opposing effects: either trauma victims are consumed with the event, and, in turn, the event consumes their sense of self, by making them unsure of who they are, what they stand for and what the world is like, or they disassociate from the event, presenting with other symptoms in the psychiatrist's office (Kirmayer, 1996). Aupperle et al. (2012) show that the cognitive deficits experienced by those with PTSD do not stop at attentional networks; trauma also affects inhibitory networks in the brain, which allow individuals to regulate their actions and thoughts, causing individuals with PTSD to engage in risk-taking behaviors.

The result is that those living with PTSD are often acting and reacting from a limited memory repertoire, a memory repertoire formed by the traumatic experience.[21] As this jar of memories takes up more space in one's mind, it crowds out the significance of memories before it. It thwarts the production of imagining alternative futures because the primary material one can

---

[21] It is important to note here that we all have an ability to restrict our memory repertoire when it suits us, particularly in instances of desiring objects we know are not good for us or others. The problem for the person who feels stuck in time in the way I have fleshed out is that they cannot revert to their fuller memory repertoire. My suggestion is that their limited memory repertoire causes their own desires to be blunted, because they lack this ability to switch back to a fuller repertoire when they please.

use to create the imaginal scenes is the material from the traumatic memory repertoire. In this way, the traumatic event draws a deep fog around the past before the trauma, and the future after the trauma, making itself seem to be the only event worthy of the individual's undivided attention. Feeling stuck in time is thus a feeling that one's mental space is no longer under one's control, even though one is trying desperately to control it. As Miller's metaphor demonstrates, one actively *locks* the memory away to forget. This has the opposite effect and results in the memory kept in the jar overflowing and restricting the capacity for MTT beyond the traumatic memory repertoire. This makes simulating different past and future episodes reliably and accurately extremely difficult.

As an attempt to overcome these attentional deficits, one often resorts to physical and emotional numbing. This is captured powerfully by Van der Kolk's (2015) research on how the body holds on to traumatic experiences long after the experience itself. Ordinarily, when individuals are not paying attention to the past or future, they are paying attention to their own states of being. They have self-awareness in this way. Individual with PTSD, on the other hand, show

> … almost no activation of any of the self-sensing areas of the brain: The MPFC [the medial prefrontal cortex], the anterior cingulate, the parietal cortex, and the insula did not light up at all; the only area that showed a slight activation was the posterior cingulate, which is responsible for basic orientation in space. There could be only one explanation for such results: In response to the trauma itself, and in coping with the dread that persisted long afterward, these patients had learned to shut down the brain areas that transmit the visceral feelings and emotions that accompany and define terror. Yet in everyday life, those same brain areas are responsible for registering the entire range of emotions and sensations that form

> the foundation of our self-awareness, our sense of who we are. What we witnessed here was a tragic adaptation: In an effort to shut off terrifying sensations, they also deadened their capacity to feel fully alive.
>
> (Van der Kolk, 2015, p. 107)

The IRP does not currently give attention an important role in MTT, either as a precondition for successful instances of MTT or as a factor enabling one to judge one's reliability and accuracy. It also is yet to consider how one's attentional resources can be eaten up by one's social situation, as in instances of trauma. Both of these are, in my view, oversights.

Narrative

Theorists in feminist philosophy, philosophy of psychiatry, and philosophy of memory converge on the idea that narrative matters to forming a coherent sense of self: forming one's own experiences into a narrative is a form of self-constitution and meaning-making (Kirmayer, 1996) and sharing such narratives helps us to understand the world around us (Brison, 2002; Fivush & Graci, 2017). Narrative theorists propose that self-narratives "function as implicit organizing structures through which we interpret and make sense of our past histories; project ourselves into the future via plans and intentions; and make sense of our actions, emotions, desires, beliefs, character traits and relations to others" (Mackenzie & Poltera, 2010, p. 32).

Trauma interferes with one's narrative. It does so by effectively splitting the self: the self before the traumatic event and the self after the event. The self which tries to reconcile the two is left in a state of disarray because the world often doesn’t make sense after traumatic experiences. Charlotte Delbo summarizes this effect recalling her nightmares of Auschwitz:

> The skin covering the memory of Auschwitz… is tough: Sometimes, however, it bursts and gives back its contents. In a dream, the will is powerless. And in these dreams, there I see myself again…scarcely able to stand…pierced with cold, filthy, gaunt, and the pain is so unbearable… that I feel it again physically, I feel it again through my whole body which becomes a block of pain, and I feel death seizing me, I feel myself die. Fortunately, in my anguish, I cry out. The cry awakens me, and I emerge from the nightmare, exhausted. It takes days for everything to return to normal, for memory to be “refilled” and for the skin of memory to mend itself. I become myself again, the one you know, who can speak to you of Auschwitz without showing any sign of distress or emotion.
>
> (Charlotte Delbo, in Kirmayer 1996, p. 186)

The question of who one is after a traumatic experience has troubled the literature on trauma for decades, as accounts of trauma express “the profound effect of living with memories that subvert the everyday construction of the self” (Kirmayer, 1996, p. 182). Traumatic memories subvert the everyday construction of the self because they make the individual feel fundamentally unsafe in both their bodies and their relations with those around them (Van der Kolk, 2015). There is a sense in which the individual living with trauma feels like they have *outlived* themselves insofar as the individual they were pre-trauma no longer makes sense to them (Brison, 2002). The traumatized individual is thus imprisoned in their traumatic memory repertoire: they know the world isn’t as they once thought it was, but dwelling on this could rob one of one's entire world. Due to the attentional deficits trauma can bring, those living with trauma can feel unable to transcend their traumatized self to recall the self before it.[22]

The IRP acknowledges that MTT is important to be able to constitute a temporal narrative of the self, but it fails to acknowledge how narrative is relationally and socially mediated. Many traumatic events are social in nature and affect marginalized groups in particular. An essential ingredient to MTT thereby depends on one's social milieu and one's relations with others. We cannot build a coherent narrative if this narrative is socially or relationally undermined.

Semantic Memory

Semantic memory has been established as an essential component of one's ability to engage in MTT. This kind of memory is "typically conceptualized as an individual's repository of acquired knowledge of the world around him or her" (Irish, 2016, p. 389). Whilst certain conditions of dementia can erode this capacity, so too can trauma. Trauma disrupts one's semantic resources because it puts an individual in a state of confusion, not only as to who they

---

[22] Strawson suggests that narrative theorists are committed to two claims, one descriptive and the other ethical. The descriptive claim is that humans generally view their lives as a story, or a collection of stories. The ethical claim is that one needs to have this ability to live a flourishing life (Strawson, 2004, 428). Strawson suggests that the ethical thesis is harmful because individuals do not need a narrative to flourish in life. Consequently, it may not alarm philosophers like Strawson that an individual with trauma feels as if their self is split. Strawson's argument depends upon a distinction he draws between those who view themselves as 'diachronic', meaning that the individual sees themselves as a temporal being with a past, present, and future, and 'episodics', who do not think of themselves this way. According to Strawson, one could live a fulfilling life as an episodic, as what truly matters to one's sense of self is how one feels and acts in the present, not the narrative one has about experiences in the past and future. However, trauma instructs us that whether one or not has the cognitive flexibility to engage in MTT bears quite precisely on how one feels and acts in the present and how one flourishes in the present moment. Sharing Strawson's sentiment, I am not preoccupied with the narrative arc of my entire life, as I recognize I am not the exact same person I was yesterday. Yet I can only realize this in virtue of my ability to engage in MTT. I only know who I was yesterday upon mentally recalling it. Without an ability to select memories, one does suffer. For Strawson's account of episodics to be successful, therefore, he must distinguish between those 'episodics' who feel they can and do still use MTT and those who cannot. Mentally ill individuals are those who often struggle to engage in MTT and ought to be recognized as a kind of episodic which cannot flourish in the ways Strawson purports they can.

are and what they remember but also as to what they *know*. The traumatized person often feels as though their entire mental library has come crashing at their feet, and they have no means or way of reorganizing the books after the event, as explained by my therapist at the time. Feeling stuck in time is thus experientially disorientating because it chronically disrupts one's worldview; it makes one's worldview so small that one can only see books of one's trauma on the floor – the one set of books one wishes they did not have to see. This disruption of semantic memory should be distinguished from disruption to attention because it concerns what one knows about the world around one. One's knowledge of the world before the traumatic event comes into question because one's ability to sort one's current knowledge is compromised. Trauma thus results in, "a fundamental reorganization of the way mind and brain manage perceptions. It changes not only how we think and what we think about, but also our very capacity to think" (Van der Kolk, 2015, p. 32).

The IRP does not presently consider that semantic memory can be disrupted by experiences such as trauma, even though the literature in psychiatry is rich with evidence that trauma does in fact disrupt this faculty of memory. This is an oversight of the IRP because if one is unsure about what they know, generating reliable and accurate future possibilities and past memories seems impossible. If one can do it, and one is only working with material from the traumatic memory repertoire, one's past will look and feel bleak, and one's future will too.

Imaginal Possibilities

Feminist philosophers have long been preoccupied with how imagination is thwarted for women in patriarchal societies (Lloyd, 2000; Mackenzie, 2000). Mackenzie in particular points out that women are often constrained in their ability to imagine themselves differently because the social world restricts women's possibilities by telling them what they ought to be and how they ought to act. One way it does this is by creating both negative and positive stereotypes, which are then internalized.[23] These limitations can affect the capacity for MTT, for instance when women are blind to opportunities they were once told they could not achieve. *In theory*, women can apply to the same jobs as men and, *in theory,* ought to be paid equally for those same jobs. Yet, in practice, women seldom apply to the same jobs as men and seldom receive equal pay when they do (Fluchtmann et al., 2021). This sense of obscured possibilities affects one's ability to MTT by restricting the future possibilities one can envision as being possible for oneself. I propose that the ways in which the social world is constituted and enacted can cause individuals to feel stuck in time because they cannot move past the constrained imaginal possibilities of the past. According to this proposal, women and other marginalized groups can have their capacity to MTT obscured by both the social imaginary and others in their milieu.

Trauma shapes imaginal possibilities in a similar way to that of marginalization: by heavily constraining possibilities. It impoverishes one's epistemic resources to envision different kinds of futures because it often limits an individual's attention to a negative memory repertoire, which thwarts one's hopes for a different kind of future (Brison, 2002), and thus results in one imagining a minimal range of future possibilities, those that are based on their traumatic repertoire. Take, for instance, how individuals with trauma avoid triggering stimuli. I avoided going for walks even though previously I had loved walks, simply to avoid the possibility of encountering a similar event. This occurs because "[b]eing traumatized means continuing to organize your life as if the trauma were still going on—unchanged and immutable—as every

---

[23] Take, for instance, the effect of beauty standards (that women must be slim, bronzed, young, and firm to be desirable) on women's perception of themselves (Widdows, 2018). Widdows argues that these qualities have become understood as an ethical standard that women must adhere to in order to be deemed 'good enough'. She argues that we see this in the language used around women who fail to meet these standards: she let herself go, she is not making the best of herself. The social punishment and judgement of women's bodies in turn produces a set of moral duties concerning beauty which women feel they must engage in. Such beauty standards are demanding and oppressive to women (Garcia 2018). This causes a sense of myopic possibilities for those who internalize these stereotypes. They are myopic in the sense that the social expectation obscures what one thinks one ought to do.

new encounter or event is contaminated by the past" (Van der Kolk, 2015, p.64). One's ability to engage in MTT is affected by this because the reliability and accuracy of one's episodic memory system are disturbed. On every walk, I would predict I would encounter something I could not cope with, something I could not bear to see. Of course, such predictions were flawed, but they were flawed because my episodic memory system was still on high alert to the relevant trauma stimuli.

At the extreme, traumatic experiences are often noted to be so unexpected and to have disturbed one's sense of imaginal possibilities so much that they experience a sense of the world as being, in one or more senses, unbearable. As one participant in a recent study on trauma and homelessness poignantly puts it: "[t]hings are traumatic if…] they are unbearable; it is beyond the imagination of the person to think that such a [thing can happen]. It is an unexpected and unbearable experience" (Gilmoor et al., 2020, p. 458). This exclamation should show that trauma itself shakes one's imaginal possibilities in a destructive way, which is likely to result in impairments in one's ability to engage in MTT.

At present, the IRP does not account for how one's social situation could affect one's ability to remember the past and imagine different kinds of futures. In so doing, it implies that each individual has the same amount and kinds of futures open to them. In principle, this ought to be how the world is, and yet, as shown, in reality, marginalized groups often have their imaginal possibilities constrained.

Cognitive Flexibility

Recall that Suddendorf and Corballis (1997) suggest that one needs to be able to disassociate from the present moment to successfully engage in MTT according to the IRP, which requires a certain cognitive flexibility. The case of trauma problematizes the idea that disassociating from the present moment is sufficient for this cognitive flexibility. Individuals with trauma, after all, tend to be disassociated from the present moment, given that their attention is firmly on the past traumatic event. Yet, they lack the cognitive flexibility required to MTT successfully because they are stuck, and their movement in time is constrained. This suggests that simply being able to disassociate from the present moment is not sufficient to indicate that one has the cognitive flexibility to engage in successful MTT. Cognitive flexibility matters to engage successfully in MTT but we cannot judge this by looking at how well one disassociates from a particular time (present, past, future, imagination) if one is consumed by, and *stuck* in, another time. This suggests that the IRP must refine its precondition for cognitive flexibility.

Disruptions to Self-Trust

The IRP overlooks the idea that self-regarding attitudes such as self-trust may play a role in MTT. Nonetheless, self-regarding attitudes appear just as critical to MTT as they are to autonomy (Dillon, 1997; Govier, 1993; Mackenzie & Stoljar, 2000). This is because, as Michaelian emphasizes, reliability and accuracy are integral to successful MTT. If this is so, one needs to be able to *judge* one’s reliability and one’s accuracy.

Trauma grossly disrupts one's ability to judge this and to have trust in oneself. Self-trust is characterized as one’s ability to cope with the world (Govier, 1993). It has both a cognitive and affective component: one must have faith in one’s competency in the area in question and one should have a generally positive attitude toward oneself to qualify as having self-trust. Trauma impairs both. Concerning the cognitive aspect of self-trust, trauma causes one to be fundamentally unsure of oneself, one’s capabilities, safety in the world, trust in one’s body to react, one’s relations, and one’s general knowledge of the world (Van der Kolk, 2015). As previously noted, the *unthinkable* and *unbearable* elements of trauma may widen an individual's sense of negative possibilities to such an extent that to use them in acts of MTT would result in only negative projections into one's past or one's future.

One domain in which the affective element of self-trust becomes particularly clear is that of mental illness. Mental illness generally erodes one's ability to have a positive self-conception. This is wonderfully put by Elyn Saks, who was diagnosed with schizophrenia:

> I was such a horrible disappointment to myself. How could I not be a disappointment to them as well…*I am a bad person, I deserve to suffer. People are talking about me. Look at them; they're staring at me.* They're talking about me. In all likelihood, that part, at least, wasn't paranoia. Given my appearance, it seems quite likely that people were talking about me.
>
> (Saks, 2007, p. 75)

Saks' account of mental illness shows that mental illness amplifies negative perceptions of oneself and others whilst dulling positive perceptions of oneself and others. This is significant as our inner imaginal conversations use data from the social world to run – we cannot imagine what we believe to be impossible (Suddendorf & Corballis, 1997). Mental illness, in a sense, opens the doors to generating negative imaginal possibilities of ourselves and others.

I argue that similarly to cases of mental illness, traumatic experiences can open up the same negative range of imaginal possibilities, particularly if one's trauma results in PTSD. Furthermore, traumatic experiences open up negative possibilities about what one can expect from the world around oneself. The culmination of these can result in individuals with trauma suffering from degradations to both the affective and cognitive elements of self-trust. In so doing, it makes it difficult for the individual to judge their reliability, as self-trust appears to be a pre-condition for being able to do this. This account thereby suggests that it is an oversight by the IRP to neglect the importance of attitudes such as self-trust as potential preconditions for one's ability to MTT successfully.

Summary of Upshot for the IRP of MTT

Examining the case of trauma highlights that the IRP of MTT is currently incomplete as it stands. There are factors the IRP considers as important to one's ability to MTT, such as narrative and semantic memory, but the IRP neglects to account for the ways in which they are socially and relationally mediated. It pinpoints the wrong locus of the precondition for cognitive flexibility, by focusing on an individual's ability to escape the present moment. So doing neglects that cognitive flexibility requires more than simply being able to transcend the present moment. Individuals with trauma can do this, yet they still suffer from impairments in agency and their ability to MTT. Lastly, there are important elements to MTT that the IRP neglects to consider, such as the role of attention, imaginal possibilities, and disruptions to self-trust, all of which seems to be preconditions for being able to successfully MTT.

In contrast to the IRP, the relational picture of MTT I defend in the following section emphasizes that all five conditions, which are a mix of affective and cognitive elements to MTT, are relationally mediated. It thereby demonstrates that one's capacity for MTT is not simply inside one's own head – others affect our capacity to be able to engage in successful MTT by either enhancing or thwarting our abilities in these areas. It suggests that the value of

MTT cannot be reduced to planning and reasoning alone, as traumatic experiences may not disrupt an individual's ability to plan or reason, but can nevertheless undermine their agency in significant ways.

§5 MTT Revised: Relational, Affective, and Cognitive Dimensions

The purpose of §4 was to highlight that MTT has a number of dimensions that can be impaired by trauma leading to the phenomenon of being stuck in time. Many of these dimensions are affective and relational. Certainly, MTT involves cognitive dimensions, as the IRP suggests. However, the fact that trauma undermines these additional dimensions of MTT indicates that it has social aspects and is heavily socially and relationally mediated. In other words, the social world both potentially supports and interferes with our ability to successfully simulate past and future possibilities. This occurs with each factor listed in §4, but most apparently in the case of imaginal possibilities and disruptions to self-trust.

By presenting the phenomenology of feeling stuck in time, I have emphasized the role of affective and emotional elements in MTT, such as one's capacity for self-trust, sense of imaginal possibilities, and narrative. Feeling stuck in time can impact these affective and emotional elements by blunting one's sense of who one is, and what one is competent to do. I have suggested that cognitive dimensions of attention, semantic memory, and flexibility are important in allowing one to successfully engage in MTT. However, I have argued that even these cognitive dimensions of MTT have social and relational aspects. The social world can divert and corrupt one's attention, semantic memory, and cognitive flexibility in significant ways. Trauma shows us how and is therefore a philosophically stimulating case to explore in reference to MTT because it enables us to defend a relational picture of what MTT is and why it matters.

Furthermore, I have argued that by examining the case of trauma, we expose several deficiencies in the IRP of MTT, both in accounts of what it is and why it matters. Trauma highlights that one's capacity for MTT is not all-or-nothing: one's ability to MTT can be impaired in degrees, and it can exist in degrees given the magnitude of dimensions involved in successful instances of MTT. What separates someone from having a fully functioning capacity for MTT and a malfunctioning one may be as simple as encountering a traumatic experience. Yet, this disruption does not rob individuals of their ability to MTT altogether, though it may well affect their ability to *successfully* MTT. By offering the precise ways in which trauma disrupts one's ability to MTT, I aimed to give a richer picture of the kinds of cognitive, affective, and relational elements which are essential to successful MTT. It also emphasized that MTT is important for more reasons than reasoning and planning. Individuals with trauma may retain an ability to plan, yet they may lack other essential elements I listed in §4.

The upshot of the new relational picture is this: if MTT is socially and relationally mediated, as I have claimed, it can be threatened by ubiquitous social and relational encounters. This is a novel suggestion in the literature and one which demands greater philosophical attention given its clinical consequences.

Chapter 1 References

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders [5th Edition],* Arlington, VA: American Psychiatric Publishing.

Aupperle, R. L., Melrose, A. J., Stein, M. B., & Paulus, M. P. (2012). Executive Function and PTSD: Disengaging From Trauma, *Neuropharmacology*, 62 (2), 686–694.

Breslau, N., Davis, G. C., & Peterson, E. (1991). Traumatic Events and Posttraumatic Stress Disorder in an Urban Population of Young Adults. *Archives of General Psychiatry,* 48 (3), 216–222.

Brison, S. J. (2002). *Aftermath: Violence and the Remaking of a Self.* USA: Princeton University Press.

Campbell, R., Dworkin, E. R., & Cabral, G. (2009). An Ecological Model of the Impact of Sexual Assault on Women's Mental Health. *Trauma, Violence, & Abuse*, 10 (3), 225–246.

Damasio, A. R., Tranel, D., & Damasio, H. (1991). Somatic Markers and the Guidance of Behaviour: Theory and Preliminary Testing. In (Eds.) A. L. Benton, H. S. Levin, & H.M. Elsenberg, *Frontal Lobe Function and Dysfunction*, 342–350. Oxford, UK: Oxford University Press.

Dillon, R. (1997). Self-Respect: Moral, Emotional, Political. *Ethics*, 107 (2), 226–249.

Fivush, R., & Graci, M. (2017). Memory and Social Identity. In (Eds.) S. Bernecker & K. Michaelian, *The Routledge Handbook of Philosophy of Memory*. UK: Routledge.

Fluchtmann, J., Glenny, A. M., Harmon, N., & Maibom, J. (2021). The Gender Application Gap: Do Men and Women Apply for the Same Jobs? *IZA Discussion Papers*, 14906.

Gerrans, P. (2018). Painful Memories. In (Eds.) D. Debus, D. Perrin, & K. Michaelian, *New Directions in Philosophy of Memory*. UK: Taylor & Francis.

Gerrans, P., & Kennett, J. (2010). Neurosentimentalism and Moral Agency. *Mind*, *119 (*475).

Gerrans, P., & Kennett, J. (2017). Mental Time Travel, Dynamic Evaluation, and Moral Agency. *Mind*, *126* (501), 259–268.

Gerrans, P., & Sander, D. (2014). Feeling the Future: Prospects for a Theory of Implicit Prospection. *Biology & Philosophy*, 29 (5), 699–710.

Gilmoor, A., Vallath, S., Regeer, B., & Bunders, J. (2020). "If somebody could just understand what I am going through, it would make all the difference": Conceptualizatons of Trauma in Homeless Populations Experiencing Severe Mental Illness. *Transcultural Psychiatry*, *57*(3), 455–467.

Govier, T. (1993). Self-Trust, Autonomy, and Self Esteem. *Hypatia*, 8 (1), 99–120.

Kennett, J. (2009). Mental Disorder, Moral Agency, and the Self. In (Eds.) B. Steinbock, *The Oxford Handbook of Bioethics*. Oxford: Oxford University Press.

Kirmayer, L. (1996) *Landscapes of Memory: Trauma, Narrative, and Dissociation,* in (Eds.) P. Antze & M. Lambek, *Tense Past: Cultural Essays in Trauma and Memory*, 173-199, Routledge: London.

Kirmayer, L., Felipe, A., & Walsh, E. (2022). The Social Determinants of Health in Neuroscience: An Overview. *Transcultural Psychiatry*, *forthcoming*.

Lloyd, G. (2000). Individuals, Responsibility, and the Philosophical Imagination. In (Eds.) N. Stoljar & C. Mackenzie, *Relational Autonomy*. Oxford: Oxford University Press.

Mackenzie, C. (2000). Imagining Oneself Otherwise. In (Eds.) N. Stoljar & C. Mackenzie, *Relational Autonomy*. Oxford: Oxford University Press.

Mackenzie, C., and Stoljar. N (2000) *Relational Autonomy: Feminist Perspectives on Autonomy, Agency and the Social Self*, Oxford: Oxford University Press.

Mackenzie, C., & Poltera, J. (2010). Narrative Integration, Fragmented Selves, and Autonomy. *Hypatia*, *25*(1), 31-54.

McLeod, C, and S. Sherwin. (2000). Relational Autonomy, Self-Trust, and Health Care for Patients Who Are Oppressed. In (Eds.) N. Stoljar and C. Mackenzie, *Relational Autonomy*, 259–279, Oxford: Oxford University Press.

Michaelian, K. (2016). *Mental Time Travel*. Cambridge, MA: MIT Press.

Michaelian, K. (2020). Confabulating as Unreliable Imagining: In Defence of the Simulationist Account of Unsuccessful Remembering. *Topoi*, 39 (1), 33–148.

Michaelian, K. (2021). Imagining the Past Reliably and Unreliably: Towards a Virtue Theory of Memory. *Synthese*, 199 (3–4), 7477–7507.

Michaelian, K., & Sutton, J. (2017). *Memory*. The Stanford Encyclopaedia of *Philosophy. Retrieved on 10th June, 2022 from https://plato.stanford.edu/archives/sum2017/entries/memory/>*

Michaelian, K., & Sutton, J. (2019). Collective Mental Time Travel: Remembering the Past and Imagining the Future Together. *Synthese*, *196*, 4933–4960.

Miller, C. (2020). *Know My Name; A Memoir*. London: Penguin.

Perrin, D., & Michaelian, K. (2017). Memory as Mental Time Travel. In (Eds.) S. Bernecker & K. Michaelian, *The Routledge Handbook of Philosophy of Memory*. UK: Routledge.

Saks, E. R. (2007). *The Center Cannot Hold: My Journey Through Madness*. New York, USA: Hyperion

Strawson, G. (2004). Against Narrativity. *Ratio, 17*(4), 428–452.

Suddendorf, T., & Corballis, M. C. (1997). Mental Time Travel and the Evolution of the Human Mind. *Genetic, Social, and General Psychology Monographs*, *123*(2), 133–167.

Suddendorf, T., & Corballis, M. C. (2007). The Evolution of Foresight: What is Mental Time Travel and is it Unique to Humans? *Behavioural and Brain Sciences*, 30 (3), 299-313.

Suddendorf, T., Addis, D. R., & Corballis, M. C. (2009). Mental Time Travel & the Shaping of the Human Mind. *Philosophical Transactions of the Royal Society, 364*, 1317–1324.

Tulving, E. (1983). *Elements of Episodic Memory*. Oxford, UK: Clarendon Press.

Van der Kolk, B. A. (2015). *The Body Keeps the Score: Brain, Mind and Body in the Healing of Trauma*. UK: Penguin Books.

Widdows, H. (2018). *Perfect Me Beauty as an Ethical Ideal.* New Jersey, USA: Princeton University Press.

Chapter 2: Memory, Colonialism, & Psychiatry – How Collective Memories Underwrite Madness

Chapter 1 introduced the idea that there can be psychological threats to MTT which can arise from ubiquitous traumatic encounters. Chapter 2 expands on this, by showing how marginalization can affect MTT in a direct way. By means of an exploration of the psychiatric ideas of Frantz Fanon, I show that racialization, the process by which racial hierarchies are made and enacted, affects both episodic memories, those we have of our personal pasts, and collective memories, the intersubjective narratives we share in our societies.

I argue that dominant collective memories ought to be rejected as a means of self-preservation for racialized individuals because these memories do not do justice to the violence of colonialism. Nevertheless, rejection of dominant collective memories comes at a significant personal cost in our societies, as it creates a traumatic loop for racialized individuals. I propose that psychiatry itself plays an instrumental role in the alienation experienced by racialized individuals because psychiatry has not yet appreciated the ways in which colonialism continues to have a hold on memory. Piecing together a theory of remembering in Fanon's texts, this article suggests that this cycle of alienation can be broken, if psychiatrists incorporate Fanon's insights regarding memory into their practice.

This chapter thereby covers cycle (ii) outlined in the introduction of the thesis, by showing how marginalization can result in impairments to MTT which may lead to one being potentially deemed "mad".

Remembering is forward-looking.
Memory is, in fact, the human capacity that equips us for the future.

Sue Campbell

(1956-2011)

§1 Introduction

Memory plays a pivotal role in our understanding of the world and in our understanding of ourselves, as our daily lives are shaped by what we remember. It also plays a fundamental role in shaping both our self-identity and our sense of belonging in our social milieu. Memory, then, seems to constitute a metaphorical landscape, charged with affective and social significance (Kirmayer, 1996, p. 175). This metaphorical landscape involves memories we have of ourselves and others, memories others have of us, collective memories of a shared past, and the needs and demands of the future. Memory is thus inherently socially constituted; we share memories with those in our social circles and this is "one of the many ways in which we come to know others and form relationships with them" (Campbell, 2015, p. 2). This dynamic conception of memory moves away from notions of memory that view memory as simply a reproductive "storehouse", a mental filing cabinet where we store memories in our mind and recall them when necessary or prompted. Whilst memory certainly can be reproductive in this way, developments in interdisciplinary memory studies and cognitive neuroscience favor instead an account of memory which emphasizes the creative processes we all use when remembering. These accounts emphasize that "remembering is…a reconstructive process, in which components of previous experience are extracted and recombined in a flexible manner" (Michaelian & Sutton, 2017).

Despite the influential role memory plays in our lives, connections between memory and mental illness continue to be undertheorized. This article seeks to flesh out one connection. It explores how colonialism - the brutal subjugation of people, lands, and cultures - infiltrates this reconstructive memory process, and it examines the alienating consequences of this for racialized individuals. §2 and §3 identify how colonialism continues to affect both episodic memories and collective memories.[24] I draw upon the psychiatric works of Frantz Fanon in '*Alienation & Freedom', 'Black Skin, White Masks',* and *'Wretched of the Earth',* as Fanon's work as a psychiatrist in both colonized Algeria and Tunisia provides compelling insights into how colonialism and memory can affect mental illness. Fanon's enriching phenomenological account of racialization proposes that one tool colonialism uses to sustain dominance in society is that of producing collective memories which serve to entrap racialized individuals. Racialization is the process by which difference and hierarchies between races are made (Al-Saji, 2013, p. 4). According to Fanon, this process involves various mechanisms, such as the projection of undesirable qualities onto others as a means of sustaining white dominance in society, as well as the creation and maintenance of certain images, ideals, and stereotypes which are intended to glorify whiteness whilst simultaneously demonizing Blackness (Al-Saji, 2013, p. 3-4). In societies with a colonial past and present, these stereotypes and ideas can be acquired through shared collective memories.

Fanon argues that collective memories underpinned by colonial structures are pernicious as this kind of memory is resistant to change (Drabinski, 2013). Drabinski highlights that for Fanon, history is a "prison" (Drabinski, 2013, p. 15). Fanon views history in this negative way

---

[24] Episodic memories refer to the kinds of memories we build as individual's regarding our own pasts. This definition will be expanded on in §2. Collective memories will be used in this article to refer to as "intersubjectively shared interpretations of a poignant common past with a high degree of affect" (Langenbacher & Shain, 2010). This definition will be fully unpacked in §3.

because history is dominated by collective memories narrated and configured by Western nations, who frequently and deliberately neglect the trauma and violence inflicted by colonialism. Drabinski contrasts this with how Fanon views memory. Fanon views memory as a potentially powerful force for political change, if memories can be re-written and re-imagined. The problem with respect to collective memories for Fanon lies with the fact that they are resistant to being changed, re-told, and re-imagined. I argue in §4 that, accordingly, for reasons of self-preservation, racialized individuals should reject collective memories underwritten by colonialism.[25] This is not to say that racialized individuals ought to reject *all* collective memories, as this would not be possible, nor desirable. To capture this, in this article I call the collective memories underpinned by colonial structures *dominant* collective memories. It is these which should be rejected. In addition, it is important to note that whereas my argument is being presented as a paradigm case of why dominant collective memories can be problematic for racialized individuals, I acknowledge that racialized individuals can, and inevitably will, differ in all kinds of context-specific and history-dependent ways from the case I present here. I also acknowledge that disconnecting from dominant collective memories is not a straightforward, linear, process. This process will look different for every individual and there may be other ways of coping, living, and managing dominant collective memories. I do not commit to exploring that task here, but I see such a project as worthwhile.

Rejecting dominant collective memories, even as a means of self-preservation, comes at a high cost. Utilizing Fanon's phenomenological account, §5 outlines some of the negative personal

---

[25] Although Fanon's conclusions apply to colonized individuals in particular, I suggest that Fanon's account can be applied to racialized groups in our society who have been impacted by the violent past of colonialism. One may acknowledge that the same conclusions Fanon draws here applies to other socially disadvantaged groups who have suffered traumatic pasts, for instance, the Jewish community. Using Fanon's account to explore these issues would require a development of how racialization interacts with sex, gender, religious affiliation etc. I do not commit to that task here, but I nevertheless think this could be done and would be a fruitful avenue for further philosophical exploration.

consequences of rejecting dominant collective memories. Specifically, I hypothesize that individuals who reject these memories may be at higher risk of being diagnosed with a mental illness in the racialized societies we live in. Individuals who suffer from mental illness have often been subject to persecution, socio-cultural violence, and segregation (Foucault, 1988, p. 245). In Foucault's words, "we often place people we deem to be "mad" in situations where 'he will be in debate with himself and his surroundings…far from being protected, he will be kept in perpetual anxiety' (Foucault, 1988, p. 245)." "Madness" then becomes a prison one can be put in at any time if one is deemed socially deviant, and which can cause one to feel severely alienated (Foucault, 1988). §5 argues that racialized individuals who disconnect from dominant collective memories can themselves be imprisoned in a similar way for appearing 'deviant'. In line with Foucault, I argue that racialized individuals therefore often feel not only alienated but also stuck in circumstances in which they are constantly in debate with themselves, due to rejecting dominant collective memories.

An undesirable consequence of this for racialized individuals is that they may come to doubt their own memory and their own narratives. This situation is even more severe for racialized individuals who find themselves in need of urgent psychiatric assistance. In these circumstances, individuals who have chosen to disavow dominant collective memories not only need to fit their memory to those of the psychiatrists but also must adapt their identity to the psychiatric institution they find themselves in in order to have a chance of getting out. I propose in §6 that this double alienation faced by racialized individuals could result in a disruption to their sense of self unless psychiatry accommodates and gives weight to the memories racialized individuals have created for themselves.

§7 presents Fanon's account of remembering as one solution to how we break this cycle of alienation in psychiatry. Fanon's theory of remembering has three main tenets. The first is that memories are relational; as our memories of the past are shaped by those around us. Consequently, it is not sufficient to simply disavow dominant collective memories of one's society as a means of self-preservation – one must be allowed to replace them with empowering collective memories which connect the individual to others. The second tenet is that memory reconstruction can benefit from starting from what we cannot remember, as recognizing where our memory fails us could help us to strengthen it. Relatedly, the final tenet of the theory is that memories of the past continue to have damaging effects on embodied affect, movement, and physiology. According to Fanon, memory reconstruction involves critically examining our bodies, as a means of both remembering the past and envisioning a new future. This article concludes by suggesting that Fanon's account of remembering can put an end to the cycle of alienation facing racialized individuals who are in need of psychiatric assistance. This article suggests that implementing theories such as Fanon's theory of remembering into psychiatry is critical to counter the unequal balance of power between patients and practitioners. It is my hope that those who do not share my political agenda will still be struck by the subtle ways memory plagues psychiatry and will be sympathetic to the idea that ethical psychiatric practice needs to give moral weight to patient's memories, particularly those the patient has created for themselves as a means of coping in an unjust world, after reading this article.

§2 Episodic Memory and Colonialism

Episodic memory is the kind of memory that "provides access to the personally experienced event, rather than just the knowledge extracted from the event" (Suddendorf & Corballis, 2007, p. 301). This kind of memory involves "reconstructing *particularities* of specific events that happened to the individual" (Suddendorf & Corballis, 2007, p. 301). This reconstructive process involves using our imaginal capacities when we remember events, and so departs from traditional theories of memory which view memory as simply reproductive. Instead, episodic memory, with its ability to reconstruct the past, "provides raw materials from which to construct and imagine possible futures" (Suddendorf & Corballis, 2007, p. 302).

Fanon's account of racialization suggests that the imaginal space used in constructing episodic memories for colonized and colonial subjects is not a space free from the influence of colonialism. This is because the imaginal space is dominated by stereotypes found in the collective unconscious.[26] Fanon's phenomenological account recalls his own feelings of objectification as he runs into images of the Black man in advertisements which sustain the negative stereotype of Black men as friendly, but unintelligent, such as *Y a bon Banania.*[27] For members of racialized groups, then, we can imagine that their imaginal space continues to be heavily shaped by racial stereotypes and thus comes to affect how they think of the past, and also what they conceive to be possible in the future.

Just as one's imaginal space is affected by the stereotypes one finds in it, Fanon's account implies that so too are the narratives we have of ourselves and others. Episodic memory involves the generation of content by utilizing past narratives to predict future outcomes. Humans "affectively broadcast" past emotions into their future, whereby individuals come

---

[26] Fanon describes the collective unconscious as functioning as a kind of racial imaginary, which helps to keep racialized subjects in their place by locking them in certain stereotypes (Al-Saji, 2013, p. 4). This collective unconscious is "neither innate nor pre-given but rather cultural and acquired" (Al-Saji, 2013, p. 4). It affects subjects as early as childhood and infiltrates subject's minds through a variety of mediums, though Fanon focuses predominantly on media, films, social engagement, stories and images (Al-Saji, 2013, p. 4).

[27] This French advertisement is used to advertise a popular chocolate drink in France and was exceptionally popular, despite its racist stereotyping and connotations. The branding itself portrays a Black man in a degrading and pejorative way, by insinuating that he can only speak simple French, thereby reinforcing negative and harmful stereotypes regarding Black people.

to predict what their experiences of others will be like as a result of past experiences and emotions (Suddendorf & Corballis, 2007). Thus, if one is hurt in the past, one will project hurt into one's future.

Significantly, we cannot expect that narratives of colonial trauma will change over generations given current research documenting how intergenerational episodic memories work. Research suggests that repressed trauma which is passed down on a generational level can produce conversion reactions in children, causing children to show distressing medically unexplained symptoms. For instance, it has been shown that, if a mother experiences repressed trauma, this affects both her world-view and how she treats her child as a result, which can lead to the child displaying unexplained symptoms (Kozlowska, 2007). Kozlowska explains that if one's parents were mistreated this could result in a world-view whereby the child expects similar treatment (Kozlowska, 2007). The child develops coping mechanisms to deal with this, which are then passed down to their offspring, who develop their own way of coping with the trauma.

That colonialism affects one's episodic memory need not by itself disrupt the self-identity of the individual. I hypothesize that distortions and disruptions of the self only occur for racialized individuals if they cannot align their episodic memory to fit dominant collective memories which have attempted to deliberately and perniciously erase knowledge regarding the atrocities of the past.

§3 Collective Memory and Colonialism

The term "collective memories" is often used to refer to a concept which is "little understood, with numerous overlapping, conflicting, or often unrelated definitions" (Boyer & Wertsch, 2009; Isurin, 2017). Maurice Halbwachs first used the term "collective memory" to identify the phenomenon of social groups that form their own memories of an event (Halbwachs, 1980). Take, for instance, family memories. Halbwachs suggested that, no matter the time at which one enters a family, customs and memories pre-existed the time of entry; these pre-existing customs and memories enable groups such as families to continue to produce their own memories of events. Halbwachs observed that individuals belong to a range of different social groups and so "there are as many collective memories as there are groups and institutions in society" (Halbwachs, 1992, p. 22). The definition of collective memories as being formed by social groups has been called the 'strong version' of collective memory (Isurin, 2017, p. 12). This is contrasted to the 'distributed version' of collective memory put forward by Bartlett (1967), which argues that, given the complexity of group social dynamics, it makes sense to move from talking about memories *of* groups, to memories *in* groups (Bartlett, 1967; Isurin, 2017). Bartlett (1967) thus rejected Halbwachs claim that groups can have memories of their own, due to diverging memories within the social group itself, but instead believed that social organization of memory, "gives a persistent framework into which all detailed recall must fit, and it very powerfully influences both the manner and the matter of recall" (Bartlett, 1967, p. 296).

In addition to the distinction between strong and distributed accounts of collective memories, there is also a distinction drawn between competitive versus multi-directional accounts of collective memories. Competitive accounts of collective memories argue that some versions of collective memories deserve to hold more moral weight than others, typically arguing that one social group's suffering is more deserving of attention than another. Multi-directional accounts of collective memories have been proposed to try to avoid pitting victims of violent hate crimes against one another. These accounts try to draw links between collective memories, rather than comparisons between moral harms committed. Multi-directional accounts suggest that collective memories are so hard to define precisely because they blur all boundaries: between individuals, collectives, social groups, and political spheres. These accounts highlight that “the borders of memory and identity are jagged; what looks at first like my own property often turns out to be borrowing or adaptation from a history that initially might seem foreign or distant” (Rothberg, 2009, p.5). They suggest that collective memories are not only porous, without clear boundaries, but span across generations, continents, and national contexts (Rothberg, 2009). Yet, collective memories are also dynamic in nature, involving processes of forgetting and remembering, and of writing and re-writing, which are subject to inner tensions and personal revisions.

For the purposes of this article, I am interested in multi-directional, distributed accounts of collective memory, which take collective memory to refer to ““intersubjectively shared interpretations of a poignant common past with a high degree of affect”, which influence “‘who a group is by providing answers of “where” it came from and “why”” (Langenbacher et al., 2010). Accepting this account of collective memories entails that within any given community, there are competing, multi-layered, and contested collective memories. Given the multi-directionality of memory, and the competing nature of social groups, not all members of a given social community will converge with respect to their collective memories of an event. The distributed, multi-directional model of collective memory suggests that this is to be expected and does not necessarily threaten the validity of the memories in question. A useful illustration is put forward by Isurin (2017) in examining the case of how Russians remember WWII differently. Isurin observes that some Russians feel a sense of pride when remembering WWII, as many Russians believe themselves to be the sole winners of the war, whereas other Russians feel an overwhelming sense of guilt when indulging in certain meals because of the memory that a million people died of starvation during the 900-day blockade of Leningrad (Isurin, 2017). Isurin claims that the reason for the difference in collective memories within social groups is that "[e]very collective memory is going to be remembered and recalled differently by individuals who might have a different personal experience of relating to such memories" (Isurin, 2017, p. 13).

Nevertheless, collective memories can be operationalized by colonial narratives so that they appear unitary, fixed, and belonging to all citizens. Collective memories which are formed in this way tend to be centered on singular and exclusionary narratives. Part of the political violence inflicted by these memories is that the narrative heralds it as the only one, thereby eradicating and silencing divergent collective memories within a given community. Novick voices this concern, claiming that, "[c]ollective memory simplifies, sees events from a single, committed perspective; is impatient with ambiguities of any kind; reduces events to mythic archetypes" (Novick, 1999, p. 4). Whilst this erasure of alternative perspectives is never complete, and even dominant collective memories rarely end up being unitary and all encompassing, I argue in §4 that collective memories of this kind are insidious and should be rejected.

Despite the fact that dominant collective memories can be problematic, for the reasons mentioned above, collective memories continue to serve important functions in our societies. Collective memories function as cultural creations that enable individuals to explain "who [they] are today and what they stand for, thereby generating emotional bonds, solidarity, and trust" (Langenbacher et al., 2010). They help individuals to have a unified social identity across time and serve as a means of relationality between people. Collective memories also play a role in the construction of shared emotions. Take, for instance, Holocaust remembrance events.

Children who may lack knowledge of the Holocaust shed tears when adults do regarding the same event. Hirsch calls these memories “post-memories”, memories that precede one’s birth which can nevertheless be traumatic to those who develop them. Children develop these memories because the original traumatic experience has been transmitted to them so powerfully by familial members that they imaginatively engage with the memory as if it was one of their own (Hirsch, 2012). She explains that these memories can be as overwhelming to those who experience them, irrespective of the distance one has from the original traumatic experience. This is often described in the literature as a process of “emotional contagion”, which occurs by people being easily affected by another’s emotional state, so much so that they come to share it, even if they originally had distance from it (León et al., 2019). Poignantly, this aspect of emotional contagion can take place without words and from exposure to certain environments alone.

Collective memories are internalized in much the same way that episodic memories are. Just as episodic memories influence content generation, so too do collective memories. They do so by influencing who is seen as belonging to a particular social collective and who isn’t. If someone is socially ‘othered’ in the collective memory, whereby they are marked out as deviant, they may be at risk at evaluating this othering into different social interactions, which can put a significant strain on relationship building. Fanon’s own account describes a poignant example of such social othering. This instance begins with a child on a train pointing at Fanon, exclaiming, “Look! A Negro!” (Fanon, 1967a, p. 89*)*. The reader sees that Fanon is made to feel utterly depersonalized by treatment that is a result of the stereotypes which frequently nest themselves into dominant collective memories. The symbolism of the child also indicates that dominant collective memories can be internalized by even innocent subjects, signifying that they are internalized in much the same way that episodic memories are. Moreover, because painful memories such as these are internalized, we are far more likely to rehearse and evaluate our options based on interactions like this.

Memory construction relies on data from the past to predict the future and make those predictions as accurate as possible. Dominant collective memories again play a part in this, by creating certain kinds of people and by controlling what is seen as relevant suffering and what is not. One means that colonialism uses to deny the reality of racialized members of society is that of controlling what counts as a collective memory. As noted previously, what is distinctive about collective memories is how historically and socially different actors are painted as either being morally good or morally bad in the societies in which we live. These expectations are then fixed by dominant collective memories, so much so that no matter what one does one cannot escape the pressure of living up to these expectations.

The upshot of this consideration is that racialized members of societies with a colonial past are often unable to merge dominant collective memories of their society with their own episodic memory, which potentially could result in a disruption in their self-identity. This disruption has consequences for psychiatry which will be explored in §6 of this article, after first demonstrating (§4) why it is that racialized individuals have good reasons to reject dominant collective memories for means of self-preservation.

§4 Rejecting Dominant Collective Memories for Self-Preservation

We have established in §3 that collective memories serve three functions: creating a unified social identity over time, relationality, and generating shared emotions. Fanon’s account suggests that for racialized people, collective memories, if underwritten by colonialism, act to achieve the opposite of all these ends.

With respect to creating a unified sense of identity over time, typically the episodic memory of members of colonized groups will be shaped by the violent past of colonialism. However, dominant collective memories of colonial societies typically elude responsibility for this violence. This can act to undermine and contradict the lived experience of racialized individuals, causing them deep intrapersonal alienation. Fanon expresses this, “I slip into corners and my long antennae pick up the catch-phrases strewn over the surface of things” (Fanon, 1967a, p. 116). Here Fanon recalls suffering from depersonalization as a result of racialization. The imagery Fanon uses here is that of feeling like an insect. Insects are vulnerable in nature, particularly to suffering violence from other creatures, and are also likely to be seen as dis-individuated, unlike mammals and humans, and are therefore not afforded the same personal differences. Nevertheless, there is a positive symbolism invoked here too by Fanon, as insects are viewed in several cultures as having transformational powers. Yet, the reader feels that Fanon is experiencing a negative, rather than a positive, transformation, as he continues, “I slip into corners, I remain silent, I strive for anonymity, for invisibility” (Fanon, 1967a, p. 116). That Fanon chooses terms like ‘anonymity’ and ‘invisibility’ is not coincidental – as the reader feels that Fanon is suffering from forgetting his own identity. Dominant collective memories underwritten by colonialism ultimately *cause* his own episodic memory to fail him.

Regarding relationality, Fanon’s account demonstrates that dominant collective memories serve as a barrier to interpersonal communication between racialized people. They do so through eradication of other languages and by upholding the importance of a dominant language and a dominant culture (Fanon, 1967a, pp. 13-16*)*. Fanon uses the example of Antilleans, who are forced to forget their own culture by adopting a new language (Fanon, 1967a, pp. 13-16). He employs a folklore story to illustrate:

> After several months of living in France, a country boy returns to his family. Noticing a farm implement, he asks his father, an old don't-pull-that-thing-on-me-peasant, "Tell me, what does one call this apparatus?" His father replies by dropping the tool on the boy's feet and the amnesia vanishes. Remarkable therapy.
>
> (Fanon, 1967a, p. 13)

Fanon's use of the term 'amnesia' here is suggestive that racialized subjects can have entire memories erased by colonialism until a physical, or mental, trauma brings back the original memory back to the surface.[28] The passage also indicates the transformative nature of returning home once one has adopted a new language. It illustrates that if one abandons one's own dialect and can no longer engage with familial members as a result, one is bound to suffer emotional pain. The racialized subject is put in a position of disconnection by adopting a new language, as they are forced to forget their own culture as a consequence of this. The family adopt a critical attitude toward the child as they no longer feel a sense of connection with their child. As one's health depends on the connections one shares with others, particularly one's relationships to one's own family (Walker & Rosen, 2005), this is no trivial discovery in Fanon's account.

---

[28] Fanon's remarks regarding amnesia are pertinent to the discussion presented so far. Amnesia can result both in memory loss of particular events but can also result in an inability to form new memories. Whilst we may initially think states of amnesia are undesirable, researchers believe that often times amnesia occurs as a survival coping mechanism (Freyd, 1996). For instance, sufferers of childhood abuse will often forget memories of such abuse, until something brings the original memory back to them (Freyd, 1996).

Dominant collective memories can also act as a barrier to generating shared emotions because they function to misrecognize the value and contribution of racialized individuals in our societies. These memories pollute our social spheres and breakdown opportunities for emotion-sharing, by causing racialized individuals to struggle for social recognition. Struggling for social recognition occurs when one is denied a social status one is worthy of (Rashed, 2019). This struggle is an emotionally negative one, whereby one feels as though they are “a person whose self-conceptions, reasons for action, and beliefs are some way invalid” (Rashed, 2019, p. 105). If one’s social setting makes one feel “invalid” in this way, through the stereotypes perpetuated in dominant collective memories, possibilities for positive emotion sharing collapse.

In summary, Fanon’s account gives us three good reasons as to why racialized individuals ought to reject dominant collective memories. The first is that, if they do not, racialized individuals will struggle to have a unified sense of self across time. The disruption caused by trying to unify their episodic memory and dominant collective memories will result in a loss of personal connectivity. The second reason, the loss of familial connectivity, follows from this. Fanon’s account suggests that the cost of endorsing dominant collective memories is nothing short of self-sacrifice and self-mutilation for racialized subjects. Fanon continues, “The immobility to which the colonized subject is condemned can be challenged only if he decides to put an end to the history of colonization” (Fanon, 1961, p. 15). Finally, Fanon’s account establishes that collective memories underwritten by colonialism arouse negative emotions for racialized individuals because they cause a social harm: the social harm of misrecognition (Rashed, 2019). To avoid these harms, Fanon’s account recommends that racialized subjects disavow dominant collective memories in Western societies as a means of self-preservation.

A potential criticism of this analysis is that memory ought to aim at the truth and that certain collective memories could be harmful but nonetheless accurate.[29] In such cases, it appears to be an epistemic vice to turn away from collective memories. Viewing memory as a moral landscape and a moral practice is the key to answering this concern (Lambek, 1996). As suggested in §1, I do not think that racialized individuals should turn away from *all* collective memories, only dominant ones underpinned by colonialism. These kinds of memories are pernicious because they often suggest that they are the only one's worth having. However, caring about 'truth' plausibly refers to being concerned with issues of accuracy and integrity (Campbell, 2014). Accuracy requires us to be concerned with selection, and integrity with how our memories not only affect ourselves but others around us. Both of these values are compatible with the idea that at any given time there is a multiplicity of viewpoints on a particular event, and both urge us to act with caution regarding dominant collective memories which cause harm to others in our communities, precisely because they may bear little relation to 'truth'. Dominant collective memories, after all, are self-selecting, aiming to eradicate alternative narratives in favor of their own, and showing little concern for how they impact the lives of those marginalized by those narratives. Take, for instance, the British portrayal of the Bengal famine of 1943 as a fault of Indian "over-breeding" rather than of British neglect. As Winston Churchill is heralded a hero in dominant collective memories in the UK, other concerns with his policies and actions are silenced. Campbell captures the way to remedy this by reiterating that "More faithful memory might require bringing one's memories into relation with different aspects of experience at that time, with what else is vivid to memory and what only vaguely recalled, and with what one now understands about that time" (Campbell, 2014, p. 43). Thus, we have strong reason to believe that caring about the truth requires listening to

---

[29] I thank an anonymous reviewer for raising this objection.

multiple perspectives on the same historical event, not silencing these perspectives, as dominant collective memories require.

§5 Risks of Rejecting Dominant Collective Memories

Fanon's account indicates that racialized subjects need to reject dominant collective memories as a means of self-preservation. This section will explore the potential negative personal costs one has to endure in doing so.

§5.1 The Memory Paradox, Racialization, & Psychiatry

> Where they find no room
> where you leave no room
> where there is absolutely no room for them
> and you dare to tell me it doesn't concern you!
> that it's no fault of yours!
>
> (Fanon, 1967b, p. 14)

Deficits in memory are thought to feature in various psychiatric conditions, ranging from dementia through to major depressive disorder (Pittenger, 2013). Thus, testing memory is of importance to psychiatrists, particularly during diagnostic assessments such as the mental status examination (MSE). The psychiatric interview typically involves reviewing the patient's history and asking the patient ad hoc questions of cultural competence (Carlat, 2017). Questions of cultural competence involve naming dates of cultural significance. Such questions may strike us as innocuous, but these questions are normative. They give significance to certain collectively-defined cultural events over others, and, in so doing, arguably undermine "the resources others need to re-experience their past in ways that meet their present needs and interests, including those of challenging dominant views of the past" (Campbell et al., 2014, p. 72). For individuals who have rejected dominant collective memories for self-preservation, questions such as these could trigger a whole range of negative feelings, producing an affective landscape which is not conducive to building rapport between the practitioner and the patient.

Questions of cultural competence aside, the patient is still expected to delve into their history and memories with their practitioner during diagnostic assessments in psychiatry. Fanon described this as psychiatrists needing to prescribe "situational diagnoses" to their patients. Situational diagnoses are those that employ considerations such as a patient's relationship with themselves, with other people, their preoccupations and occupation, their sexuality, their sense of security and insecurity, and their life narrative as part of diagnosis and the prescription of treatment (Fanon, 1976b, p. 10). All of this standardly occurs by asking the patient questions during the assessment interview. Fanon mocked this process with respect to how racialized patients are treated by psychiatrists, exclaiming, "Is there not something a little comical about speaking of the North African's relations with his associates, in France? Does he *have* relations? Does he *have* associates? Is he not alone?" (Fanon, 1967b, p. 10). Fanon's writings at the time were situated in a particular context of colonization and so some may suggest that this conclusion cannot be applied to psychiatry today. However, it seems that what Fanon had to say about psychiatry then still affects patients today. Take for instance the case of David Harewood, a Black British actor, and his father, both of whom had experiences in psychiatric institutions in the UK. David writes:

> When he got out of the psychiatric ward, my father was a very different man…My father was bitter, angry at the system – in particular the doctors and nurses for thinking he was crazy – and he was angry at those who didn't see things the way he'd seen them.
>
> (Harewood, 2021)

Harewood elaborates that this failure to see the impact of racialization has severe mental and physical consequences for himself and his father. He recounts his first experience of racism in the UK as twisting apart his identity at the seams. Harewood recalls an instance where he was told to "Get the fuck out of my country, you little Black bastard!" by an elderly man in his youth (Harewood, 2021). He recalls this experience from memory in the following way:

> I was shocked and rooted to the spot. He glared at me, before slowly turning and walking away. I watched as he strode off, working out what he had just said to me. "His" country? Was it not mine, too? Perhaps it wasn't? I was piecing it together, my imaginary game now over, replaced by more serious thought. Maybe I don't belong here? And that's when it happened, the two halves of me split. There was now a Black half and an English half and I could feel myself slowly coming apart. At times in my life I've been able to fuse these two halves together, but occasionally the gap between them is just too big. How can I be English when much of England refuses to accept my Blackness and makes it clear that I am not welcome?
>
> (Harewood, 2021)

This "splitting" of the self could, and tragically did for David, result in mental distress and affective breakdown. Yet, the standard set of psychiatric interview questions do not ask whether patients have experienced distress as a result of racialization (Carlat, 2017). They instead inquire about sets of relationships, such as friendships, family relations, and sexual relations. Such questions do not ask about the relationships one has in general with others in their social milieu. Memories of racialization are left aside during diagnostic assessments.

They are also left aside once an individual is institutionalized. That memories are viewed as important during diagnostic assessments, but then dismissed as unimportant when a patient is institutionalized, is the “memory paradox” in psychiatry. Researchers Brown and Reavey acknowledge this, highlighting:

> On speaking to patients and staff [at the forensic psychiatric unit], it became clear that references to the past – to significant events, to recollections of time before admission – were not deemed relevant to the smooth running of the ward, which was focused exclusively on the medical management of the patient.
>
> (Brown & Reavey, 2015, p. 156)

As noted above, David Harewood’s father expressed bitterness and appeared a different man after being institutionalized for a breakdown caused by racist abuse in the UK. If the abuse which caused David Harewood’s father to have a breakdown is not acknowledged in either the psychiatric assessment nor institutionalization – there is no room for the patient to disclose or discuss their experiences of racialization and their unique coping methods in response to its effects on their psyche – it can hardly be a surprise that racialized patients are bitter towards the system. One potential negative consequence of suffocating one’s experiences to fit the system is that the patient may come to doubt the self they have created outside the system. This could cause racialized patients to feel more “splits” in their sense of self, the opposite of what psychiatry and psychiatrists want to achieve. I suggest that this creates a traumatic loop for racialized individuals, particularly those in need of urgent psychiatric assistance, whereby their memories and narratives of the past are doubted so much that they come to doubt their own memories, as well as their sense of self. This traumatic loop involves feeling unwell due to racist abuse or the effects racialization has on the psyche, then being unable to voice these concerns to the practitioners who desire to help you. Not having the opportunity to voice these concerns can make patients feel as though these memories are not important, thereby invalidating their own negative experiences of racialization. This can have effects on the self which are explored in §6.

§5.2 Psychiatric Institutions and Alienation

The patient-practitioner relationship is not the only tool of alienation psychiatry has at its disposal. The psychiatric institution itself poses a risk of alienating individuals by imposing a given identity on them. Foucault’s critique of mental asylums first captured the concern, by identifying how cultural myths, most of which are shaped by dominant collective memories, work to sustain barbaric treatment of the mentally unwell in the psychiatric institution. He summarized, “beneath the myths themselves, there was an operation, or rather a series of operations, which silently organized the world of the asylum, the methods of cure, and at the same time the concrete experience of madness” (Foucault, 1988). According to Foucault, psychiatric institutions have a history of perpetuating cycles of madness through alienation.

Erving Goffman’s ethnographic work helped to explore how this institutional alienation occurs. He suggested that psychiatric institutions are “total institutions”, which determine the identity of the patients that are in them (Goffman, 1991). “Total institutions” refer to institutions which in some sense sever individuals ties with their own world, thereby creating an entirely new world in which the individual is forced to live. Goffman argues that this severance enables psychiatric institutions, and other institutions like prison, to determine the patient's identity and disrupt their sense of self (Goffman, 1991). Goffman suggests that the patient's new identity, once constructed by the institution, must align with the ambitions of the psychiatrist.

These findings have been advanced more recently by Suman Fernando (Fernando, 2017). Fernando notes that the very history of psychiatry shows us that the need to diagnose mental illness is tied in a very direct way with the need to control certain populations. Fernando argues that the mental health system is an institution set up as a means of controlling certain racialized groups in various settings. His work argues that racialized people in the United Kingdom are at higher risk of the following: being diagnosed as schizophrenic, being sectioned under the Mental Health Act, being admitted as offender patients, being transferred from open wards to closed ones, and being held by the police under the Mental Health Act investigation (Fernando, 2017). We may not initially perceive these issues as being issues concerning memory. However, dominant collective memories are a colonialist tool which are at play here because it is these memories which suggest that certain populations "need" controlling in the first place.

One salient conclusion to be drawn from this body of psychiatric and ethnographic work is that psychiatric institutions can constitute patients' realities and, in turn, manufacture their identities if they erase patient voices and memories for purposes of institutionalization. A tool they utilize to do so which is often overlooked is by controlling and manipulating memory. As alluded to in §5 once a patient is institutionalized, they are often treated as though their past is irrelevant, and, painfully, as if their future is too. Describing a patient in the forensic psychiatric unit who confessed feeling sadness about losing her sexuality on the ward, researchers Brown and Reavey describe her as being, "a body without a past and, apparently, with little sense of the future" (Brown & Reavey, 2015). To ignore a patient's past is to ignore the patient's future. Nonetheless, this is exactly how "wellness" is performed in psychiatric institutions: "it does not require reference to the patient's past experiences before they entered the unit" (Brown & Reavey, 2015, p. 161). When patients rebel against the systematized control of their memories, by reporting past experiences, not only are these experiences not always taken seriously, they can be undermined. This concern is captured by a patient who states:

> Sometimes you tell them things [the staff] and they don't give a shit about it [the patient here is referring to an alleged sexual assault by a staff member]. It's probably like being on trial. You have told one of the doctors, the junior doctors, and she just said to me, *I can't really do anything about it because it was in the past, in the past and we are here now.*
>
> (emphasis mine, Brown & Reavey, 2015, p. 166)

Patients in psychiatric institutions are thus often apprehended in the present for their "wellness", yet, psychiatrists often ignore "how that present is being continuously pushed and pulled by forces from the outside" (Brown & Reavey, 2015, p. 20). The ways in which patients are constructed inside psychiatric institutions are directly tied to the ways in which they are constructed outside these institutions. Thus, the identity imposed by the psychiatric institution will not only be shaped by dominant collective memories of the given society but also poses the risk of causing racialized individuals to feel personally alienated by needing to meet the expectations of the institution.

§5.3 Disavowal as Dysfunction

Psychiatric diagnoses have been used in the past to describe healthy behavior as disordered as a means of controlling certain populations. Take for instance the condition drapetomania, which was used to describe slaves who ran away from captivity, as well as the medicalization of homosexuality as recently as the 1980s. These cases serve as a reminder that psychiatric conditions can be invented to repress socially undesirable behavior. Rejecting dominant collective memories may well be viewed as socially undesirable behavior in Western societies which desire to glorify such memories. In 2020, Black Lives Matter protesters tore down the statue of slave trader Edward Colston in Bristol and measures were taken to protect the monument of Winston Churchill in preparation of further protests. The protesters were labelled by the Prime Minister as engaging in an act of "lying about our history" (Giordano, 2020). Boris Johnson exclaimed, "We cannot now try to edit or censor our past" (Giordano, 2020). One of the gravest harms inflicted by dominant collective memories is that re-imagining them can become demonized in precisely this way. Johnson failed to appreciate that the protesters were simply trying to re-evaluate the socio-political significance of past events. This recent example illustrates the potential condemnation of individuals who fail to endorse dominant collective memories and who seek to right past wrongs by advocating for a different form of collective memory, whilst also showing the empowerment which can arise if individuals are allowed to seek alternate forms of collective memory, outside of the chains of dominant collective memories. My suggestion is that this adverse reaction to the disavowal of dominant collective memories is likely to affect psychiatry itself.

In the past, the only members of society who could not make political demands were the "mad" (Foucault, 1988; Rashed, 2019). Hence, we call those whom we wish to silence "mad": whether that be women, racialized individuals who reject subordination, political figures who challenge the status quo, individuals who reject entrenched gender norms, or individuals who reject boundaries on whom they can love and when. Rashed notes, “once madness is delegated to physicians and conceived of as mental illness, no further dialogue is possible” (Rashed, 2019) and this explains why pushing back on the medical monologue about madness is so important to activists (Rashed, 2019). Given that socio-cultural expectations and political demands desire to silence certain narratives and often harness the power held by psychiatrists to do so, I hypothesize that one undesirable personal consequence racialized individuals face for disavowing dominant collective memories is to be at higher risk of being diagnosed with a mental disorder.

Transcultural psychiatry originated to address issues which arise from cultural difference and to overcome the issues raised in conventional psychiatry mentioned in §5.1 and §5.2. Whilst transcultural psychiatry does provide a framework to address some of the issues I have raised in §5, it does not go far enough. Specifically, one cannot ignore that transcultural psychiatry itself has battled with how best to address these issues (Kirmayer, 2012). When cultural difference shrinks, the psychiatrist may become blind to the ways that cultural differences impact the patient’s lived experience and the therapeutic relationship itself. This problem is articulated by psychiatrist Samuel Okpaku (Okpaku, 1998, p. xxi) who notes, “my supervisor asked, and I paraphrase, “Is this family different from your family?” I responded, “Not really”. With much clinical experience behind me now… My revised answer is “yes and no,” and therein lies the complexity of transcultural psychiatry.” In the case of racialized individuals who do not present as ‘culturally different’ to the psychiatrist because the difference itself, not endorsing dominant collective memories, is not perceived as salient enough, a problem arises: the problem of obscured cultural difference, the problem of “yes and no”. Rejecting dominant collective memories seems a good candidate as qualifying as one of these “yes and no” cases, the most difficult for psychiatrists to spot and, in turn, one of the cases which may slip through the net of transcultural psychiatry.

§6 Disruption to the Self: Psychiatry, Memory, & Colonialism

I have proposed that disavowing dominant collective memories gives racialized individuals freedom to build a coherent sense of self for themselves. As §5 illustrates, the problem raised by both diagnostic practices in psychiatry and psychiatric institutions themselves is that they threaten to not only undermine this created sense of self for racialized individuals, but also call into question whether the sense of self created by the individual is 'normal'. This can further alienate racialized individuals in need of psychiatric assistance. Such alienation could potentially produce a disruption in a person's sense of self.

This suggestion is significant because it offers a working hypothesis as to why it is that places like the United Kingdom, with a strong colonial past, have high rates of schizophrenia in racialized individuals (Fernando, 2017). This occurs because the self constructed by racialized individuals is rejected on an institutional level, leaving racialized individuals to feel both deeply alienated and confused. As Al-Saji notes, "If racism is reflected not only in economic, social, and political conditions, but also structures lived experience itself, then anomalies and breakdowns in experience cannot be studied as purely individual afflictions in racial societies" (Al-Saji, 2013, p. 2). Indeed, one positive consequence of my account is to show that experiencing disruptions in the self in the racialized societies we live in is normal and non-pathological, and ought to be treated as such in psychiatry.

Undermining one’s memory is an insidious way to challenge one’s sense of self, but it is sadly an effective one. Patients in psychiatric institutions are often not afforded the epistemic authority to verify their own memories of events, as shown in §5.2, and are therefore denied their right to rewrite their pasts and, accordingly, script their futures. Patients rely on figures such as psychiatric nurses and doctors to help them with this process. As with all power dynamics, this can be abused, particularly if one individual qualifies as a “vulnerable rememberer” who cannot be trusted and one is viewed as an epistemic authority on what qualifies as successful remembering. In order to restore balance between the practitioner and the patient with respect to memory, I argue that Fanon’s own theory of remembering should be adopted.

§7 Fanon’s Theory of Remembering

Theories of remembering typically fall into two categories: preservationist theories and generationist theories (Michaelian & Sutton, 2017). Preservationist theories of remembering typically take the act of remembering to be a matter of encoding information, storing it, and retrieving it at a later time (Draaisma, 2000). Generationist theories, on the other hand, take remembering to be an active process in which an individual more or less accurately represents events of the past (Wagoner, 2017). Generationist theories are linked to the idea that remembering one’s past is linked to imagining one’s future (Michaelian & Sutton, 2017). It follows on these accounts that people are not only capable of mental time travel, between the past and the future, but that this is genuinely how we begin to construct memories (Suddendorf & Corballis, 2007). [30]

---

[30] Mental time travel is a term coined by Suddendorf and Corballis (2007) to refer to “the faculty that allows humans to mentally project themselves backwards in time to re-live, or forwards to pre-live, events” (Suddendorf & Corballis, 1997).

Fanon's call for a new humanism both invokes memories of atrocity in the past, but also memories of a future, one which is still to be created (Fanon, 1961). Fanon seemingly foresaw generationist theories of remembering and, I argue, his own theory of remembering falls into this category. Fanon's account of remembering does not make a causal connection to an event a necessary condition for having a memory of the event in question. Rather, like modern simulation theories of memory, his account begins by suggesting that appropriate causal connection is not a prerequisite for successful remembering (Michaelian & Sutton, 2017). [31]

The process of reconstructing memories successfully, for Fanon, begins by identifying memories as *relational,* insofar as our memory of the past is shaped by the memories of those around us. This is important, as current theories of remembering, at best, underestimate this point (Michaelian & Sutton, 2017). Fanon's experimental psychiatric method, the concept of the day hospital, emphasizes the importance of relational wellness to mental health.[32] Keeping connections, Fanon argues, in particular familial ties, is crucial to the success of mental health interventions (Fanon, 1964, p. 498). Fanon describes his technique as, "No cut occurs here" and "No longer is the patient a patient with severed antennas" (Fanon, 1964, p. 501). The imagery invoked, of loss of connection equating to severing and cutting patients, indicates that disconnecting patients from their ties by isolating them in a psychiatric institution is likely to do more direct harm than good. We cannot justify psychiatric interventions which recommend this and can no longer deny patients their right to their pasts.

---

[31] Fanon's suggestion here is a powerful one, which threatens to undermine the way the DSM-V is currently set up. For instance, the current diagnostics for post-traumatic stress disorder in DSM-V do not seem to reflect the idea that one can be traumatized by colonialism even if one did not have direct family members who were violently affected by it. Fanon's account implies that this is a mistake.

[32] Fanon designed an experimental method of day hospitalization, whereby patients stay at the psychiatric institution during the day but are allowed to return home in the evening. The principles of the day hospital include: the patient is not ostracized from their family milieu or professional work unnecessarily and the psychiatric symptomology that patients present with do not disappear because of internment (Fanon, 1964, p. 474).

However, as argued for in §4, Fanon's account establishes that for racialized subjects, familial and other connections cannot be found in dominant collective memories underwritten by colonialism. As such, in order to reconstruct memories successfully, patients need to substitute dominant collective memories with alternative collective memories that the patient finds empowering instead. These alternative collective memories could be the ones shared by family members, or by other communities with which the patient feels a connection. This requires psychiatrists to acknowledge the value of memories distinct from dominant collective memories in their own society. It also requires, Fanon suggests, psychiatrists understanding the cure for mental illness, which he describes as, "the cure – the ordered calling into question of the pathological structures – must proceed precisely at the core of the syncopated dialogue established between the overall personality and its environment" (Fanon, 1964, p. 504). Fanon proposes that, contrary to the belief that the patient can only become 'well' when they are separated from their environment, that the reverse is true – the patient can only become 'well' when they are allowed to locate and vocalize injustices in their social environment and have these concerns listened to and appreciated. The potential for this dialogue can be severed due to an unhealthy power dynamic where the patient's vulnerabilities are easily exploited. Fanon calls for a new kind of medical humility and an awareness not only of one's positionality in power dynamics, but for a corresponding accountability for their positionality on the part of psychiatrists.

Fanon acknowledges that "[t]he essential weapon of psychiatry is psychotherapy, which is to say a dialogue between the mental patient and the doctor" (Fanon, 1964, p. 405*)*. Fanon warns the reader that one means psychiatry can use to undermine patients is through exploiting speech and memory. By applying Fanon's critical openness to the process of memory construction psychiatrists could avoid committing this harm. Unlike traditional theories of remembering which center on how successfully one can remember events in one's past, Fanon's psychiatric method favors the opposite approach: focus on failures and learn from them (Fanon, 1964, p. 353). Fanon acknowledges that he must adapt his methods to different kinds of patients and be humble enough to be receptive to the needs of his patients (Fanon, 1964). It is important that the psychiatrist not only engages in this practice, but that they lead by example as Fanon did on this front. Fanon's theory of remembering calls for psychiatrists to be vulnerable about their own memories and share times they have misremembered events, as well as what they cannot remember from their own pasts with their patients. This should be followed by psychiatrists addressing what they can remember. Fanon's methodology suggests that doing so is advantageous because it disrupts the power dynamic in psychiatric encounters and is useful in identifying what successful remembering looks like to the individual.

Memory reconstruction which centers on what one cannot remember also has an additional value for psychiatrists like Fanon: it gives patients the freedom to acknowledge, in light of the psychiatrist's declaration, that failures of memory do not equate to a failing episodic memory. This breaks the traumatic loop described in §5, as patients come to realize that there is not something intrinsically wrong with their memory, and, correspondingly, the self they have created.

The final piece of memory reconstruction for Fanon lies with the body. According to Fanon, trauma does not simply affect the mind; it also affects the body. This leads to physical bodily pains (Fanon, 1967b) and the development of psychosomatic disorders (Fanon, 1967a). Fanon describes the past as "burning" and continues, "He does not understand that anyone should wish to impose on him, even by way of memory, the pain that is already gone" (Fanon, 1967b,

p. 2). Fanon's suggestion that an individual's inability to face the past may result in bodily pain in the present is significant, given the mounting body of contemporary research connecting memory and the body (Mingon & Sutton, 2021; Narváez, 2013; Yehuda & Lehrner, 2018). Accordingly, we can, and ought to, expect racialized bodies to remember the trauma of the past and hold memories of this trauma in their bodies. This gives us reason to suspect that racialized bodies may exhibit similar bodily reactions to the psychosomatic disorders examined by Fanon or to expect that these memories may continue to plague bodies in ways we have not yet properly examined in psychiatry. By placing the body as one of the core pillars of memory reconstruction, Fanon's account breaks the traumatic loop mentioned in §5, because it suggests that irrespective of the patient's mental state, the patient remains a knower with respect to memory: the patient knows how they feel in their bodies and, even more importantly, can share this knowledge and the roots of these feelings with their psychiatrist.

In addition, Fanon's account calls the attention of psychiatrists to the ways in which the body is made and sedimented through embodied collective memories. As Narváez asks, "For how can a society penetrate the soul of a citizen without touching her body first, without first addressing her clothes, her appearance, her tone of voice, her manners and her appetites?" (Narváez, 2013). As noted in §5, psychiatrists may not be alert to the ways that dominant collective memories themselves come to be embodied by their patients in this way. Behnke (1997) used the phrase "ghost gestures" to refer to lived micro-movements which shape how people make and re-make their bodies. She identifies that, "inner gestures of numbing oneself create the body of violence" (Behnke, 1997, p. 195) and claims that racialized individuals may be haunted by these gestures. Mukandi (2019) reflects on this point in reference to Fanon, movement, and the psychiatric encounter. He suggests that clinical encounters are stages for performance where "there are (at least) two actors, both engaged in a long running drama, whose performances demand interpretation" (Mukandi, 2019, p. 415). Mukandi argues that there does not need to be any malice or ill-intent on the part of psychiatrists to still end up reinforcing dominant collective memories, such as the one which suggests the North African client is unintelligible (Mukandi, 2019, p. 416). So doing puts an unfair burden on the racialized patient in the clinical encounter to 'move' to meet the clinician where they are with respect to their memorial landscape: "the client, the individual in pain or in distress, must first undertake the movement necessary to be intelligible to the steadfast clinician. This is the work the client that is "other" must do" (Mukandi, 2019, p. 422). Fanon's emphasis on the body aims to take seriously these claims about movement in psychiatry. Mukandi poignantly summarizes the point of Fanonian humanism as being an act psychiatrists must engage in "to traverse the distance necessary in order to reckon with the other's humanity" (Mukandi, 2019, p. 424). The power of Fanon's theory of remembering lies in giving psychiatrists tools to do precisely that.

§8 Conclusion

This article has explored the ways in which colonialism has infiltrated the reconstructive memory process and suggested possibilities for it to be positively disrupted, both on an individual level and on an institutional level in psychiatry. I have presented Fanon's account as one explanation as to why dominant collective memories should be rejected as a means of self-preservation for racialized individuals. However, rejecting collective memories comes at a significant personal cost in the racialized societies we live in, the cost of experiencing a traumatic loop for the individual who rejects them. Fanon's theory of remembering can break this loop, if taken seriously by psychiatrists.

I hope to have illustrated that what is at stake when we refuse to respond to the moral significance of past events is not simply that we are failing to be faithful to the multiple, different, equally valid ways of appreciating the past. The real harm is that on an individual and institutional level we are committing social harms towards members of our communities by failing to acknowledge the significance of the past to the present and to the collective future. There is an imperative for a culturally-informed and decolonial psychiatric project that calls for this cycle to be circumvented, by practitioners, communities, and individuals alike.

Chapter 2 References

Al-Saji, A. (2013). Too Late: Racialized Time and the Closure of the Past, *Insights*, 6 (5), 1-13.

American Psychiatric Association, (2013). *Diagnostic and Statistical Manual of Mental Disorders [5th Edition],* Arlington, VA: American Psychiatric Publishing.

Behnke, E. A. (1997) Ghost Gestures: Phenomenological Investigations of Bodily Micromovements and Their Intercorporeal Implications, *Human Studies, 20 (2), 181-201.*

Boyer, P., & Wertsch, J. (2009). *Memory in Mind and Culture.* Cambridge: Cambridge University Press.

Brown, S. D., & Reavey, P. (2015) *Vital Memory and Affect: Living with a Difficult Past.* New York: Routledge.

Campbell, S. (2014). *Our Faithfulness to the Past: The Ethics and Politics of Memory,* (Ed). *C.M. Koggel & R. Jacobsen,* Oxford: Oxford University Press.

Carlat, D. J. (2017) *The Psychiatric Interview.* Philadelphia: Wolters Kluwer.

Draaisma, D. (2000) *Metaphors of Memory: A History of Ideas About the Mind.* Cambridge: Cambridge University Press.

Drabinski, J. E. (2013). Fanon's Two Memories, *South Atlantic Quarterly,* 112 (1), 5-22.

Fanon, F. (1961). *Wretched of the Earth,* R Philcox (trans.), Grove Press: New York.

Fanon, F. (1964). *Alienation and Freedom,* Ed. J. Khalfa and R. Young, S. Corcoran (trans.), New York: Bloomsbury Press.

Fanon, F. (1967a). *Black Skin, White Masks,* C Markmann (trans.), Grove Press: New York.

Fanon, F. (1967b). *Toward the African Revolution, Political Essays,* H. Chevalier (trans.), New York: Monthly Review Press.

León, F., Szanto, T., & Zahavi, D. (2019). Emotional Sharing and the Extended Mind, *Synthese,* 196 (12), 4847 – 4867.

Fernando, S. (2017). *Institutional Racism in Psychiatry and Clinical Psychology: Race Matters in Mental Health,* Cambridge: Palgrave MacMillan.

Foucault, M. (1988). *Madness and Civilization: A History of Insanity in the Age of Reason*, UK: Vintage.

Giordano, C. (2020). *Boris Johnson Says Attacking Statues is "Lying About Our History" and Protests Have Been "Hijacked by Extremists",* The Independent, Retrieved 10th June, 2022: https://www.independent.co.uk/news/uk/politics/boris-johnson-statues-churchill-mandela-colston-protests-black-lives-matter-a9562626.html

Freyd, J. J. (1996). *Betrayal Trauma: The Logic of Forgetting Childhood Abuse,* Cambridge: Harvard University Press.

Goffman, E. (1991). *Asylums: Essays on the Social Situation of Mental Patients and Other Inmates,* London: Penguin.

Hacking, I. (1995). *Rewriting the Soul: Multiple Personality and the Sciences of Memory*, Princeton University Press: New Jersey.

Hacking, I. (1999). *The Social Construction of What?* USA: Harvard University Press.

Halbwachs, M. (1980) *The Collective Memory.* New York, NY: Harper & Row.

Halbwachs, M. (1992) *On Collective Memory,* Lewis. Coser (trans.), Chicago, IL: University of Chicago Press.

Harewood, D. (2021) *Maybe I Don't Belong Here: A Memoir of Race, Identity, Breakdown and Recovery,* Bluebird: UK.

Harewood, D. (2021). 'I came close to death': David Harewood on Racism and Psychosis, The Guardian. Retrieved on 10th June 2022, from https://www.theguardian.com/culture/2021/aug/21/david-harewood-homeland-racism-psychosis-close-to-death

Hirsch, M. (2012). The Generation of Post-Memory: Writing and Visual Culture After the Holocaust. New York: Columbia University Press.

Isurin, L. (2017). *Collective Remembering: Memory in the World and in the Mind.* Cambridge, UK: Cambridge University Press.

Kirmayer, L. (1996) *Landscapes of Memory: Trauma, Narrative, and Dissociation,* in (Ed.) P. Antze & M. Lambek *Tense Past: Cultural Essays in Trauma and Memory*, 173-199, Routledge: London.

Kirmayer, L. (2012). Rethinking Cultural Competence, *Transcultural Psychiatry*, 49 (2), 149-164.

Kozlowska, K. (2007). Intergenerational Processes, Attachment, and Unexplained Medical Symptoms, *Australian and New Zealand Journal of Family Therapy*, 28 (2), 88-99.

Langenbacher, E., Shain., Y., & Ebrary, I. (2010) *Power and the Past: Collective Memory and International Relations.* USA: Georgetown University Press.

Lambek, M. (1996). *The Past Imperfect: Remembering as Moral Practice*, in P. Antze & M. Lambek (ed.), *Tense Past: Cultural Essays in Trauma and Memory*, 235-255, Routledge: London.

Michaelian, K., & Sutton, J. (2017). *Memory.* The Stanford Encyclopaedia of *Philosophy. Retrieved on 10th June, 2022 from https://plato.stanford.edu/archives/sum2017/entries/memory/>*

Mingon, M., & Sutton, J. (2021) Why Robots Can't Haka: Skilled Performance and Embodied Knowledge in the Mãori Haka, *Synthese*, 199, 4337-4365.

Mukandi, B. (2019) *The North African Syndrome: Traversing the Distance to the Cultural "Other*" in (Ed.) S. Tekin & R. Bluhm, *Bloomsbury Companion to the Philosophy of Psychiatry*, 413-428, New York: Bloomsbury Academic.

Narváez, R. F. (2013). *Embodied Collective Memory: The Making and Unmaking of Human Nature.* USA: University Press America.

Novick, P. (1999) *The Holocaust in American Life.* USA: Mariner books.

Okpaku, S. (1998). *Clinical Methods in Transcultural Psychiatry,* USA: American Psychiatric Association Publishing.

Pittenger C. (2013). Disorders of Memory and Plasticity in Psychiatric Disease, *Dialogues Clinincal Neuroscience*, 15 (4): 455-63.

Rashed, M. A. (2019). *Madness and the Demand for Recognition: A Philosophical Inquiry into Identity and Mental Health Activism.* Oxford: Oxford University Press.

Rothberg, M. (2009). *Multidirectional Memory: Remembering the Holocaust in the Age of Decolonization.* Stanford: Stanford University Press.

Suddendorf, T., & Corballis, M. C. (2007). The Evolution of Foresight: What is Mental Time Travel and is it Unique to Humans? *Behavioral and Brain Sciences*, 30 (3), 299-313.

Wagoner, B. (2017). *The Constructive Mind: Bartlett's Psychology in Reconstruction.* Cambridge: Cambridge University Press.

Walker, M., & Rosen, W.B. (2005). *How Connections Heal: Stories from Relational-Cultural Therapy.* New York: Guilford Press.

Yehuda, R., & Lehrner, A. (2018). Intergenerational Transmission of Trauma Effects: Putative Role of Epigenetic Mechanisms, *World Psychiatry,* 17 (3), 243-257.

Chapter 3: Cognitive Transformation, Dementia, and the Moral Weight of Advance Directives

Whilst I have defended the idea that the psychological is social in Chapter 1 of this thesis, I have still spoken of psychological and social threats as if they do not converge. Yet, psychological and social threats to MTT *do* converge in extreme cases. One such extreme case which will be the focus point of Chapters 3 and 4 is that of dementia. Individuals with dementia suffer from severe memory impairments, which impact their ability to engage in MTT. They also suffer from marginalization which results from their condition, from their wider community, but also from loved ones. Chapter 3 explores one negative side to the dual threat to MTT posed by dementia: individuals with dementia often have their current preferences silenced in favor of preferences they held before they received a diagnosis of dementia.

This chapter problematizes this way of silencing individuals with dementia because I believe it neglects to account for the ways in which the individual with dementia has transformed. Dementia patients in the moderate-to-late stage of the disease can, and often do, express different preferences than they did at the onset of their condition. My proposal here is that ignoring the current preferences of an individual with dementia further marginalizes them in virtue of their memory impairments and is problematic for this reason. Chapter 4 follows on from this by showing how individuals with dementia can be positively held in relations of interpersonal trust which can help to mitigate some of the harms caused by their memory impairment.

§1 Introduction

In standard medical decision-making, patients are free to make their own choices regarding their treatment, so long as the patient in question is judged by medical professionals to have the requisite mental capacity to make an informed decision.[33] When capacity is compromised by a condition such as late-stage dementia, the patient is judged to no longer have the ability to know what is in their best interests.[34] In such cases, the responsibility for the decision of how to act in the patient's best interests is left in the hands of medical professionals or family members. Making this decision is particularly problematic, even harrowing, in cases in which the patient's current preferences contradict their preferences as expressed in an advance directive.[35] Advance directives are documents that allow individuals to set out directions and preferences for their future medical care, in the eventuality that they lose their decisional capacity (Boer et al., 2010, p. 202).[36] These documents have traditionally been conceived of as an effective means of extending an individual's autonomy from their current self, as an individual who has decisional capacity, onto their future self, who lacks it.

Advance directives currently have varying degrees of legal strength on a global scale (Vezzoni, 2005). The strongest legal status for these kinds of documents can be found in countries such as the Netherlands, Belgium, Denmark, and the United States, where advance directives are legally binding on physicians (Boer et al., 2010, pp. 202-203). Meanwhile, countries such as

---

[33] In medical contexts, determining decisional capacity is a part of every substantive interaction between a patient and their physician (Fellows 2015). Typically, determining decisional capacity involves a physician making an assessment of a patient's ability to express, make, and understand the consequences of, a particular decision. This assessment can also be done with regards to a specific task (Moye, J. & Marson, D., 2007, p. 3).

[34] Dementia has arbitrarily been divided into three dominant stages: early, mid-stage, and late-stage. Often these stages of the disease blur, but late-stage dementia is characterized by a severity of functional impairments, which cause the patient to require full-time care. In particular, the patient may find themselves easily disorientated, frustrated at their inability to do daily tasks such as clothing and washing themselves, and their perceptions may become distorted or blurred (Wells 1977, 5-7).

[35] Preferences are desires individuals have for one thing over another, which show a degree of stability and are not entirely whimsical. For instance, Popeye consistently prefers spinach rather than candy and demonstrates this by consistently choosing to eat spinach overeating candy (Fehige, C. & Wessels, U. 2011, xii).

[36] The most common kinds of advance directives are treatment and proxy directives. The former specifies what treatment, if any, the individual would desire in the future, and the latter specifies another person who is empowered to make decisions on the patient's behalf. Whilst less common, advance directives can also be written for euthanasia (Boer et al., 2010, 202).

Germany, Switzerland, and Norway give these documents ‘weak’ legal status. In circumstances such as these, whilst advance directives are viewed as having some legal status, this status does not hold decisive moral weight in medical decision-making (Boer et al., 2010, p. 203). Meanwhile, countries such as Japan give advance directives no legal status at all (Boer et al., 2010, p. 203).

The purpose of this article is to argue that the received view in the philosophical literature, which defends the strongest legal status of advance directives, is philosophically inadequate as well as out of touch with clinical practice. Clinicians show reluctance to implement an advance directive when it is in contrast with the well-being and current preferences of the dementia patient in question. Significantly, this is not simply the case in small sample of cases; both ethnographic and compliance studies indicate that clinicians have a “low opinion of the effectiveness of advance directives and are inclined not to follow treatment directives in case the content of the directive differs from, or is opposed to, their medical judgement” (Boer et al., 2010).[37] This article offers a philosophical defense of current clinical practice, by arguing that preferences made after a transformative experience such as having dementia are legitimate and ought to be given moral weight in medical decision-making. As such preference changes are unpredictable, given the nature of transformative experience itself, they could not be fully considered by someone who is in a process of drawing up an advance directive.

The structure of the article is as follows. §2 will outline the received view regarding the moral weight of advance directives, which is developed by Ronald Dworkin.[38] Dworkin’s account relies on the idea that a person’s critical interests show a degree of stability over time. I suggest

---

[37] This is further evidenced by studies by Vezzoni (2005) and The et al. (2002).

[38] This view has also been defended by other prominent scholars such as Jefferson McMahan (2005) and Govind Persad (2018).

that Dworkin's account fails to consider cases of transformative experience, where a person's preferences can change dramatically.[39] I elaborate on this consideration in §3. Here, I present the idea that undergoing a cognitive transformative experience such as dementia ought to mitigate the moral weight we assign to the patient's advance directive. This section will also consider why cognitive transformative experiences differ in kind from other transformative experiences such as giving birth for the first time. §4 considers what upshot this has for the issue of advance directives. It is not a consequence of my account that persons shouldn't be allowed to choose to give their advance directive significant moral weight if they so wish. Moreover, if a patient is aware of and has considered the implications of what I discuss in §3, and still wishes to give full moral weight to their directive, then this is permissible on my view. My account should be viewed as an argument in favor of reducing one's confidence in the effectiveness of advance directives and as an explanation as to why it is currently morally problematic to allow such documents decisive moral weight in medical decision-making. I conclude in §5 that preferences made after cognitive transformative experiences are legitimate. Clinician's beliefs regarding the limitations of advance directives suggest that they already realize this; I suggest that philosophers need to do the same.[40]

§2 Advance Directives and Dementia: The Received View

The received view in the philosophical literature, put forward by Ronald Dworkin (1994), argues that advance directives ought to be given decisive moral weight in medical decision-making. Dworkin (1994, pp. 219-235) claims that certain decisions, ones regarding the cessation and prolongation of life, ought to lie with individuals (Shiffrin, 2004). He argues that

---

[39] Transformative experiences are experiences which have the power to change your preferences, values, beliefs, or knowledge with respect to a certain domain. The notion will be fully expanded upon in §3 of the article.

[40] I thank an anonymous reviewer to bringing my awareness to this fact.

advance directives reflect an individual's judgements about their own lives and, as such, should be given significant moral weight, even if they direct medical professionals to curtail the life at stake (Shiffrin, 2004). The following real-life case studies serve to illustrate the complexity of this issue.

The first case study involves a woman called Mrs. Black.[41] At the age of eighty-five, Mrs. Black received a diagnosis of mid-stage dementia. Due to the progression of her dementia, Mrs. Black often struggled to recall the names and faces of family members. Nevertheless, she was noted by her nurses at the residential aged care facility she lived at as being an exceptionally happy woman. She took joy in her daily activities, particularly watching birds pass by through the window. Whilst in care, Mrs. Black developed a serious bacterial infection. She had an advance directive stipulating that if she were ever to suffer an illness which resulted in her inability to recognize her family members, she would not wish to receive any medical treatment to prolong her life. Consequently, her son insisted that her advance directive be followed and so the directive was implemented. She died shortly after. The entire medical team report feeling "devastated". Mrs. Black's nurse Lily surmised, "when Mrs. Black wrote the advance directive, she probably did not consider the idea that she could have dementia and still enjoy a good quality of life" (Sokolowski, 2018).

The second case is that of Mr. White, who received his diagnosis of early-stage dementia at age sixty-five.[42] As the disease rapidly progressed, Mr. White was put in residential care. Two months after being in care, he received a new diagnosis of having an inoperable brain tumor. After this diagnosis, both of his children convinced him to update his living will; at which time

---

[41] See Sokolowski (2018; p. 45-83).
[42] See Sokolowski (2018: p. 85-121).

an advance directive was drawn up for his care, naming both of his children as attorneys for his personal care. Mr. White's state declined rapidly after this. His physician explained to him, and his children, that without a feeding tube being inserted into his stomach, he would likely die soon, as a result of his inability to take food or water by mouth. Mr. White gave inconsistent answers to his physician regarding the insertion of the feeding tube. His children also had different views on whether the tube should be inserted or not. His son referred to the fact that his advance directive stated for "no heroic measures" to be taken to save his life. His son cited this as evidence that his father would not wish to have the tube inserted. His daughter, on the other hand, stated both that she was unsure that her father had the requisite capacity to write the directive and that the statement "no heroic measures" was too vague to be interpreted as meaning her father wouldn't want the feeding tube. As a result, the physician decided to ask Mr. White one last time regarding whether he wanted the tube inserted or not. Mr. White expressed a wish for the tube and so it was inserted; contrary to his son's wishes and, according to his son, contrary to the instructions set out in his advance directive.

Dworkin presents two arguments as to why he believes advance directives ought to hold decisive moral weight in cases such of these: the argument from autonomy and the argument from beneficence. [43] The argument from autonomy begins by acknowledging that competent adults have a right to autonomy, by which we mean a right to make decisions for themselves (Dworkin, 1994, p. 222).[44] Dworkin illustrates how much moral weight we give to this right, particularly in medical contexts, by referring to the example of a Jehovah's witness who has the right to refuse life-saving blood transfusions because so doing is not in keeping with their

---

[43] The right to autonomy in this context refers to the right an individual who has decisional capacity has to make decisions regarding their own medical care.

[44] Competency is a legal concept, which is typically informed by a physician's judgement of decisional capacity. Legal standards for determining competency vary, but such standards usually require that the patient is able to express a choice, to understand the relevant information in making a particular choice, to appreciate the medical consequences of their decision, and to be able to reason properly about their options (Applebaum, 2007).

religious convictions (Dworkin, 1994, p. 222). Dworkin then asks: when is that right lost? Is the right lost, or compromised, in certain cases where individuals lack certain cognitive capacities?

In order to answer this question, Dworkin suggests that we need an account of why autonomy is so fundamentally important to us in the first place. He considers two such accounts: the evidentiary view of autonomy and the integrity view of autonomy. The evidentiary view maintains that we should respect decisions individuals make for themselves, even if they are objectively undesirable, because individuals are often the best judges of what is in their best interests. Dworkin points out that if we accept this view of autonomy, we should not extend the right of autonomy to the severely demented because such persons have lost the requisite mental capacities to truly know what is in their best interests (Dworkin, 1994, p. 223).[45] Nevertheless, Dworkin rejects the evidentiary view of autonomy as persuasive because it fails to explain when and why people have the right to autonomy. He asks us to imagine that we know someone who likes to smoke and enjoys doing so even though it harms their overall well-being. Dworkin states, "If we believe…that respecting their autonomy means allowing them to act in this way, we cannot accept that the point of autonomy is to protect an agent's welfare" (Dworkin, 1994, p. 223). As a result, Dworkin argues that we cannot think that the right to autonomy ought to be respected for this reason alone, as it is not clear that individuals are the best judges of what is in their best interests. Dworkin thus rejects this view as an adequate explanation of why autonomy is important to us.

Another explanation Dworkin considers is given by the integrity view of autonomy. This view postulates that autonomy is important because it "encourages and protects people's general

---

[45] Dworkin uses the term 'severely demented' to describe those in a moderate-late stage of the disease.

capacity to lead their lives out of a distinctive sense of their own character" (Dworkin, 1994, p. 224).[46] According to this view of autonomy, people can, and often do, make choices which harm their welfare.[47] Nonetheless, it follows from the integrity view of autonomy that it is still worthwhile to allow individuals to make these kinds of personal choices for themselves. According to this view, whether or not the severely demented have a right to autonomy hinges on the degree of their general capacity to lead a life out of a sense of their character (Dworkin, 1994, p. 225). Dworkin gives us reason to think that it is unlikely that severely demented patients have such a capacity, as dementia patients in the moderate-late stage of the disease often have whimsical preferences which may well end up contradicting one another (Dworkin,1994, p. 225).[48] As such, Dworkin argues that dementia patients who are in the latter stages of the disease have lost their right to autonomy because they have lost the mental capacities necessary for the attribution of the right to make sense (Shiffrin, 2004).

Dworkin concludes that patients with severe dementia have "no right that *any* decision be respected just out of concern for their autonomy" and that this is in accordance with either view of autonomy, integrity or evidentiary, we endorse (Dworkin, 1994, p. 226). Dworkin then considers the question: what about a patient's *precedent* autonomy? (Dworkin, 1994, p. 226) He asks us to suppose that a dementia patient, called Margo, was competent in the past before developing dementia. In this time, Margo writes an advance directive which states that she should not receive life-saving treatment for any life-threatening illness she may contract (Dworkin, 1994, p. 226). He states that if we accept the evidentiary view of autonomy, we

---

[46] A good example of an individual leading a life out of a distinctive sense of their own character is illustrated by the Jehovah's witness example mentioned previously, whereby the individual refuses life-saving treatment as a result of upholding their religious conviction.

[47] This is the case both with the smoker example and also with the Jehovah's witness who refuses life-saving treatment.

[48] An example of a whimsical preference could be stating that you want a certain kind of meal preparing, only to state that you don't desire it on its arrival. In this sense, it is a fleeting preference.

wouldn’t be compelled to accept Margo’s instructions given that her preferences are likely completely changed in her new demented state. Nevertheless, Dworkin argues that if we accept the integrity view of autonomy, it follows that Margo’s past wishes should be respected because writing an advance directive is making exactly the kind of judgement that autonomy, according to the integrity view, most respects, “a judgement about the overall shape of the kind of life she wants to have led” (Dworkin, 1994, p. 226). As Dworkin endorses this latter view, this is the position he ends up holding.

Dworkin then makes a second argument as to why advance directives should hold moral weight in cases of dementia: the argument from beneficence. Dworkin argues that even though the severely demented lack rights to autonomy, they retain a right to beneficence, the right that decisions be made in their best interests. He claims that despite this, it isn’t clear what follows from this right, or what the consequences of possessing this right are. To begin to answer the question as to what is in a dementia patient’s best interests, Dworkin draws a distinction between two forms of prudential interests - experiential and critical interests (Dworkin, 1994, p. 229). Critical interests concern what is fundamentally important to the overall success of our lives. According to Dworkin, such interests are a distinct set of prudential concerns which, when fulfilled, genuinely make an person’s life better (Dworkin, 1994, p. 202).[49] Moreover, our critical interests reflect who we are and what is fundamentally important to us and our convictions about what helps to make a life good on the whole. For instance, one could have a critical interest that one’s relationships are meaningful to them and so put a lot of effort in to maintaining them. These interests are contrasted with a different kind of prudential interest: experiential interests (Dworkin, 1994, p. 226). Experiential interests are concerned with the

---

[49] Prudential interests, for the purpose of this article, are concerns which are related to well-being (Crisp, 2018).

quality of our own experiences and with pleasurable experiences, such as eating a delicious slice of chocolate cake.[50]

According to Dworkin, satisfying critical interests has priority over satisfying experiential interests. This is because the satisfaction of the former is essential to our well-being in a way the latter is not (Dworkin, 1994, pp. 230-232). Dworkin demonstrates this point by contrasting the importance of a critical interest, in this example, the critical interest of valuing your relationships with your children, "I do think my life would have been worse had I never understood the importance of being close to my children" with the effect of an experiential interest "My life is not a worse life to have lived – I have nothing to regret, less to take shame in – because I have suffered in the dentist's chair" (Dworkin, 1994, pp. 201-202). Dworkin states that severely demented patients are:

> ignorant of self – not as an amnesiac is, not simply because they cannot identify their pasts, but more fundamentally, because they have no sense of a whole life, a past joined to a future, that could be the object of any evaluation or concern as a whole. They cannot have projects or plans of the kind that leading a critical life requires. They therefore have no contemporary opinions about their own critical interests.

---

[50] We can make sense of this distinction if we are careful to pick the right examples. These examples typically contrast an important critical interest with a trivial experiential interest. My previous example of contrasting one's meaningful relationships with one's chocolate cake eating is an illustration of this. Given my example, it is reasonable to think that, in ordinary circumstances, one should favor pursuing meaningful relationships over one's chocolate cake-eating and so critical interests are more important than experiential ones. Nevertheless, the persuasiveness of this distinction can lose credibility when we alter the experiential interest in question. For instance, imagine that we changed the example of the experiential interest to one where an individual experienced acute pain. It isn't clear in this case that one would think more about one's relationships than one's chronic pain.

> (Dworkin, 1994, p. 230)

Herein lies the reason why Dworkin thinks we should implement the advance directive of the severely demented patient to promote their best interests. In line with Dworkin's conception that patients with severe dementia lack certain capacities, it follows that these patients lack the ability to have a sense of their own lives as extending over time and therefore fail to engage with, or have contemporary opinions about, the critical interests they held as a person before getting dementia. As on Dworkin's account these interests matter most when considering a patient's best interests, we should look to, and adhere to, the critical interests they expressed in their advance directive(s), even if their contemporaneous pleas contradict their advance directive.

The argument from autonomy relies on us believing, like Dworkin does, that dementia patients in the moderate-late stage of the disease lack the mental capacities necessary to lead a life out of character. Dworkin seems to make this assumption based on thinking about dementia patients who lack stability of preferences. Whilst such dementia patients do exist, it is certainly misrepresentative of a large portion of moderate-to-late stage dementia patients who do show a degree of stability in their preferences, even though these preferences may have altered drastically from the preferences the patient showed at the onset of their disease. We cannot assume that because a dementia patient's preferences drastically change over the course of their illness that this is indicative of a degradation of their autonomy. It may well be, or it could be that such preference changes are the result of undergoing a transformative experience. As we cannot know either way, I argue that we ought to take seriously the preference changes a dementia patient experiences throughout the course of their illness and give them due moral weight. This argument will be developed in §3.

In a similar vein, the argument from beneficence relies on us believing that critical interests are the more important interests and that these interests show a degree of stability and resistance to change across time. I argue in §3 that transformative experiences are unique insofar as they have the power to radically alter our critical interests. If this is true, we cannot assume, as Dworkin does, that such interests will necessarily show a degree of stability across one's lifetime.

§3 Transformative Experiences

L.A. Paul's book *Transformative Experience* introduces the idea that some perspectives are epistemically inaccessible to us until we have experienced them (Paul, 2016). For instance, a baby may not be able to know what chocolate ice-cream tastes like until they have eaten it (Paul, 2016, p. 14). Likewise, adults who wish to have children will not know what it is like to have children until they have them.

These two experiences, whilst both being transformative in some sense, are transformative in different ways. The former, the baby's ice-cream tasting, is epistemically transformative insofar as the baby's experience of eating the ice-cream for the first time will give them some new information in virtue of having the experience (Paul, 2016, p. 14).[51] The latter, choosing to have children, is personally transformative insofar as once one becomes a parent, there is a sense in which your whole being changes; the way you experience yourself changes and so too may your critical interests. Importantly, these two kinds of transformative experience are not mutually exclusive; indeed, Paul's book primarily focuses on experiences which involve both aspects (Paul, 2016, p. 17).

---

[51] The use of the word 'epistemically' here refers to gaining knowledge through the transformative experience.

Paul gives two different, but equally important, examples of kinds of experiences which involve both epistemic and personal transformative elements. The first is as follows. Imagine you are presented with the opportunity to become a vampire. With one bite, you will be changed into "an elegant and fabulous creature of the night" (Paul, 2016, p. 2). Unsurprisingly, "as a member of the undead, your life will be completely different" (Paul, 2016, p. 2). However, we cannot know exactly how this experience will feel for us until we experience it because "you cannot compare the character of the lived experience of what it is like to be you…to the character of the lived experience of what it would be like to be a vampire" (Paul, 2016, p. 4). In the second example, Paul asks us to imagine that a group of neuroscientists invent a microchip that when implanted gives humans a new sensory ability, a sixth sense in addition to the usual five (Paul, 2016, p. 6). Because the chip produces a sense we have never experienced before, Paul claims, "we cannot know what it would be like to experience the new, sixth, sense" (Paul, 2016, p. 7).

Paul presents these cases as a means of motivating the following question: given the options above how could you possibly make an informed choice? She gives us reasons to suggest that making the choice based on what you think you would like as a vampire, or as an individual with a sixth sense, is morally problematic because they involve you assigning a subjective value to an experience which is epistemically inaccessible to you from your current standpoint.[52]

§3.1 Cognitive Transformative Experiences

---

[52] Subjective values are grounded by what it is like to have certain experiences. For instance, I assign a high subjective value to my puppy-petting experience because I know I have enjoyed this experience in the past.

The kinds of transformative experiences discussed by Paul are wide-ranging. However, there is one area which her account fails to address in depth: cases of cognitive transformative experiences. A cognitive transformative experience is one which alters a person's cognitive capacities in such a way that may change the way the person thinks about their preferences, values, and beliefs. For instance, individuals who experience religious transformative experiences change not only who they are as a person, but their beliefs, values, and lifestyle. This article focuses on one particular kind of cognitive transformative experience: the experience of having dementia. Whilst this cognitive transformative experience has both epistemic and personal transformative effects, it also significantly differs to the cases Paul considers. This section will outline these differences. In spite of these differences, I argue that Paul's conclusion still holds. That is, I suggest, as Paul does, that it is morally problematic to assign a subjective value to an experience, like dementia, which is epistemically inaccessible to you from your current standpoint. This has important consequences for the issue of advance directives, which will be explored in §4. I ultimately use Paul's framework to defend the idea that preference changes which result after cognitive transformative experiences ought to be given moral weight in medical decision-making.

In order to understand the ways in which dementia differs as a transformative experience, it is necessary to begin by fleshing out what dementia is. The symptomatology of dementia is characterized by defects in several areas: memory, orientation, intellectual function and ability, judgement and affectivity (Wells, 1977, p. 2). Early-stage dementia may result in the patient reporting problems with short-term memory and also that they are "not themselves" to their physicians (Wells, 1997, p. 3). This alteration of character affects each individual differently and can be described as the patient experiencing an alteration of mood, creativity, enthusiasm, and a capacity to give and receive affection. Such alterations in character can be positive, negative, or neutral, depending on the individual whom is affected. As one can see from this description, early-stage dementia does not differentiate itself clearly from other disorders such as depression and anxiety, which can result in a person with dementia not receiving a diagnosis of the disorder until much later in the disease, where other symptoms tend to arise (Wells, 1977, p. 4).

As the dementia progresses, achievement of personal ambitions and fulfilment of social responsibilities becomes less important to the individual (Wells, 1977, p. 5). During this time, the patient becomes more likely to experience irritability and depression as their awareness of their diminished abilities becomes apparent (Wells, 1977, p. 6). In the moderate-stage, the dementia patient can experience trouble adjusting to change, following plans, and initiating activities.

In the late-stages of the disease, the patient often requires full-time care. Family members can describe their loved ones as being a “different” person. Some patients in this stage are depressed, lethargic, and lacking in energy; whereas others tend to be overreactive and enjoy social engagement once more. Patients in this stage experience a marked impairment in their learning ability, in their spatial awareness, and exhibit signs of motor and/or sensory dysfunction (Wells, 1977, pp. 6-7).

This description of dementia motivates one immediate difference between the kinds of transformative experiences Paul discusses and cognitive transformative diseases such as dementia: choice. In other instances of transformative experiences, the adult in question chooses to undergo the transformative experience. This is not the case for people diagnosed with dementia. One may argue that those who choose to have transformative experiences are, to a certain extent, liable for their decision and what follows from their decision. No such argument can be made with respect to patients who experience dementia.

A second, more substantial difference, is that dementia patients experience a gradual preference change, which is in contrast to the cases Paul presents. Unlike cases of becoming a vampire, dementia patients do not wake up a radically different kind of being. Instead, they have gradual preference revelations dependent on which stage of dementia they find themselves in. As Paul's transformative experience framework suggests, there is no way individuals can rationally assign a subjective value to an experience which is epistemically inaccessible to them from their current standpoint (Paul, 2016, p. 120). The fact that dementia patients are likely to experience many preference revelations, in a short time span, gives us a good reason to think that it is simply not possible to unify a person's values in the kind of way Dworkin seems to think is possible, as outlined in §2. If one experiences a cognitive transformative experience like dementia, one's values and desires will change in ways that one will not be able to predict or control. Significantly, trying to predict and control this experience may actually harm one's future self.

A related difference is that dementia changes not only what people come to value, but also the way in which they come to value it. In so doing, it can be argued that persons who experience dementia experience an altered sense of appreciation: both for their condition and for their prior tastes. Mrs. Black appeared to experience this change. In the early stages of dementia, she expressed that she did not wish to have the disease. Motivated by this thought, she wrote her advance directive. Nevertheless, as a dementia patient in the moderate-stage of the disease, she no longer felt that way and was markedly content. Significantly, the reverse is true in certain dementia patients.[53]

Similarly, dementia patients can, and often do, experience a difference in appreciation for their prior tastes. This is beautifully captured by the early-stage dementia patient Wendy Mitchell in her book, *Somebody that I used to know:*

> We wouldn't get on now, you and I. Too much time has passed. We are friends who have lost touch, who now lead parallel lives. We like different things. You love the work and bustle of a busy city, whereas some days I lose hours just looking out of a window at the view.
>
> (Mitchell, 2018, p. 9)

Here Wendy explains that, from her perspective, the person she was pre-dementia feels like a different person to the person she is now, who has dementia, because they have a different appreciation for the same activities. She goes on to explore this fact elsewhere in the book, saying that she used to enjoy the taste of mushrooms but no longer gets pleasure from this experience. Instead, she gets pleasure from other activities.

Nevertheless, in order to qualify as experiencing a change in appreciation, it is necessary that one's ability to apply facts to one's values changes. For instance, we can say that I change my appreciation for meat, if knowing the facts, I decide to no longer eat it for ethical reasons. What is not clear in patients with dementia is the extent to which their appreciation for certain values

---

[53] A case of this nature will be explored in §4 of the article, as these situations raise particular ethical concerns which are of importance to the discussion at hand.

changes and the extent to which this appreciation shift results as a failure to use and consider the relevant facts. As such, medical professionals often debate whether or not the preference change made by the dementia patient is truly authentic. I propose that there is an ambiguity at play here, as the dementia patient's preferences may change due to capacity *or* from the transformative nature of dementia itself.

An interesting analogy can be made here between dementia and cases of change preference during child-birth. Akin to writing an advance directive, women write down what pain-relief, if any, they want to receive during child-birth. They do so whilst they have decisional capacity because they know that their preferences will likely change once the child-birth process is underway. Once the child-birth process occurs, however, these preferences can, and often do, change. One reason for this is that until you give birth to a child, you remain in an epistemically impoverished position insofar as you cannot imagine the pain of what it will be like to give birth. Whilst one could argue that the pain of child-birth puts one in a cognitively compromised situation, insofar as you may well lack the conditions necessary for decisional capacity in such a state of pain, no one holds you to your prior commitment to not receiving pain relief if you ask for it during the process.

This is where the case of child-birth and dementia come apart. Whilst the woman in child-birth can override her prior commitment to not receiving pain relief, someone in a state of severe dementia cannot override their prior commitment to not receiving pain relief. Why is this? The difference seems to lie in that we think that the dementia patient's preferences for pain relief have changed due to a change in cognitive capacity and *not* as a result of the transformative experience. This is unlike child-birth, where we think the transformative experience of giving birth is altering the prior preference.

I propose that when a patient undergoes a cognitive transformative experience, the preference change the patient experiences may be due to diminished decisional capacity, or it may in fact arise as a result of the cognitive transformative experience itself. As we have no way of knowing what has caused the new preference to arise in the dementia patient, we ought to take their newly revealed preferences seriously if we are concerned with acting in the patient's overall best interests. This fact, in conjunction with Paul's argument that we should be cautious about constraining future choices if we do not understand what having a particular transformative experience will be like, ought to reduce our confidence in the effectiveness of advance directives in cases which involve dementia patients.

One objection to my account is that persons undergo transformative experiences all the time and this by itself is not a reason to override their prior commitments. For instance, imagine you desired to have a child and now it is born, you wish you didn't have it. One may argue that you cannot override your prior moral and legal commitment by claiming that your preferences have now changed because the child has arrived.

However, both legally and morally, preference changes experienced by parents in circumstances like these are given moral weight, even if they are frowned upon in certain circumstances. This again is dissimilar to the case of dementia, where dementia patient's preference changes are not given moral weight and allowed to override the weight of their advance directive.

In fact, there are good reasons to think that the dementia patient's preferences should be given more moral weight in comparison to the parent who changes their preferences after having their child. To begin with, whilst one can claim to lack a crucial awareness of what having a child will be like, one knows what one is morally and legally committing themselves to when having a child. It isn’t clear that, as advance directive writing stands, one is aware of the moral weight of advance directives in medical contexts. In addition, one’s agency does not unravel in cases of personally transformative experiences. In cases of dementia in particular, one moves from a certain stage of agency at time *t1* when writing the advance directive, to a different stage of agency by *t2,* when one is severely demented. For instance, one may move from a stage of personhood, where one has full cognitive capacity and awareness, to a state of selfhood, where one only has an awareness of themselves as a being over time.[54] I argue that the two cases, the transformative experience of dementia and the transformative experience of choosing to have a child, are not comparable for this reason.

Finally, one’s appreciation of their disorder may also transform after experiencing cognitive transformation. The literature on the disability paradox gives us good reason to think that this is possible, as persons who have not experienced living with disability are less likely to rate their lives with a disability as having a high level of well-being (Albrecht & Devlieger, 1999, pp. 977-988). The concept of cognitive transformative experience offers us an explanation of why this is: persons who are yet to have transformative experiences are not in a position to adequately imagine what their life will be like.[55] With respect to cognitive transformation, as

---

[54] David DeGrazia draws this distinction between three kinds of agency: personhood, selfhood, and subjecthood. one experiences a personhood kind of agency if one has the capacity to think, feel and deliberate, and has full mental capacity. This is the kind of agency most commonly attributed to adult humans (DeGrazia, 1999, p. 382). The second stage of agency is ‘selfhood’ (DeGrazia, 1999, p. 382). Selves are agents who, unlike persons, may not have higher-order functioning, but do have self-awareness over time. For instance, whilst young children may fall short of hitting the target of the kind of higher order functioning specified in the personhood criteria, we accept that children do have an awareness over time and, thus, can be classified as selves. The final stage of agency is ‘subjecthood.’ One is a subject when one has experiences and has the consciousness necessary to appreciate them (DeGrazia 1999, 382-83). For example, you could qualify as a subject if, due to a severe car crash, you had suffered severe cognitive impairments, but you were conscious and still able to feel pleasure.

[55] One alternative explanation as to why imagination fails in this way is presented by Ubel et al., (2005) who argue that people often mis-imagine what it is like to have a disability because they focus too narrowly on the most substantive difference between their life now and their life with the disability.

Lily, Mrs. Black's nurse pointed out, it is likely that people use this misguided assumption, that they will not enjoy their life with dementia, as motivation to write an advance directive. This is problematic given that we know that people can live a life with well-being and have dementia.

§4 The Upshot for the Moral Weight of Advance Directives

This section will examine how the transformative experience framework outlined in §3 should inform advance directive writing. I suggest that the framework generates three key principles individuals should keep in mind before writing an advance directive. I concede at the end of the section that there may be additional reasons for wanting to write an advance directive which are not covered in this article. I end this section by generating some tentative advice on what to do if, despite understanding that dementia is a cognitive transformative experience which one cannot predict or control, one still wants to bind oneself to an advance directive.

The first principle generated from the transformative experience framework is that one ought to be cautious about projecting current preferences onto a future self. One explanation of why this is comes from the general reflection principle proposed by Bas Van Fraassen. This principle runs as follows:

> My current opinion about event E must lie in the range spanned by the possible opinions I may come to have about E at later time t, as far as my present opinion is concerned.
>
> (Fraassen, 1995, pp. 16-17)

If someone fails to meet the principle, then the person cannot regard herself as following a rational policy for opinion change. My account has presented reasons for thinking that before you have dementia you are not in a position to reasonably imagine the possible preferences you may come to have at the time when you are demented, due to the cognitive transformation involved. It is easy to see how the disability paradox can negatively play out in advance directive writing: one assumes that one will not enjoy having dementia and this impacts the kinds of statements one writes into their advance directive. These statements can then affect a dementia patient living contentedly with the disease, as was the case with Mrs. Black mentioned in §2 of this article. The underlying motivation as to why the clinicians in Mrs. Black's case were so reluctant to abide by her directive seem to be grounded in the fact that they were highly skeptical that Mrs. Black even attempted to entertain the idea that she could live contentedly with dementia.

The second principle generated from the framework is that substituted judgement standards are problematic for patients who have dementia, precisely because they undergo a personal cognitive transformative experience. These kinds of judgement standards attempt to unify the values of the pre-dementia patient with the values of the dementia patient which is often not possible. This is because dementia changes what we value, and how we come to value it, gradually, over time. The concept of transformative experience gives us good reason to suspect that a person's values across one's whole life will not be in unison, as Dworkin and other philosophers assume that they are (see also McMahan, 2005; Persad, 2018).

One could respond that advance directives where a surrogate decision-maker is named could elude this criticism because they do not necessarily bind the dementia patient to their prior values in this way. Such measures allow for interpretation, and for the final word to go to the surrogate decision-maker. Nevertheless, this measure relies on trusting that the surrogate decision-maker will make a decision which is in the best interests of the dementia patient and not in the interests of the pre-dementia patient.[56] Both the cases of Mrs. Black and Mr. White show why this assumption is problematic. Mrs. Black's son desired that the advance directive be implemented because his mother, pre-dementia, did not want to be the kind of individual who could not remember her children. The transformative experience framework suggests that this motivation is irrational for wanting to write, and be legally bound to, an advance directive. Likewise, in the case of Mr. White, his children fought over how to interpret the advance directive in question. Thus, the conjunction of principles one and two seems to suggest that if you are motivated to write an advance directive either to project your current preferences onto a future self or as an altruistic measure to protect your family, both motivations are flawed. In the first instance because it is simply not possible to project your current preferences onto a future self in any kind of unified way. In the second instance because writing an advance directive which is ambiguous could lead to problems of interpretation and, in turn, more familial distress.[57]

The final principle is generated by considering the most rational course of action to take given the discussion so far about the problems of writing and interpreting advance directives. The transformative experience framework suggests that the rational action to take is to allow yourself to experience 'preference revelations' as the disease itself unfolds (Paul, 2016). Such

---

[56] This strategy is also problematic given what the literature on leeway in advance directives informs us. Subjects vary greatly as to how much 'leeway' they would give to their surrogate decision maker to override their advance directive. A study on advance directives and dialysis patients illustrates this: 39% of the patients desired to grant 'no leeway' to their surrogate, 19% a 'little leeway', 11% a 'lot of leeway', and 31% 'complete leeway' (Sehgal et al 1992). A patient from a different study in the leeway literature nicely captures the concern: "if you have a lot of confidence in the other person to have your best interests at heart, it would be O.K., but that may not always be the case, even with a relative" (Fried et al 2013).

[57] It is worth noting here that advance directive writing does not necessarily have to be ambiguous. I am merely suggesting that the kinds of advance directives which are currently written tend to be ambiguous and thus often lead to interpretative problems of the nature outlined here.

preference revelations can be positive or negative. A positive preference revelation may be an increased appreciation of social interactions in the later stages of having dementia, in contrast to not enjoying social interactions in the earlier stages of the disease. A negative preference revelation can be seen by the following case study of Mr. O’Connor (Jaworska, 1999). Before developing dementia, Mr. O’Connor was a deeply religious man, who believed strongly in the sanctity of life. He declared that he would always want life-saving measures to be taken in the eventuality he ever got sick. Nevertheless, Mr. O’Connor develops dementia, and shortly after loses his wife. He continually reports after the loss of his wife to others that he does not want to go on living. What does my account suggest is best to do in cases such as these, assuming that Mr. O’Connor contracts an infection which threatens his life and has an advance directive which commits him to his prior religious beliefs? Mr. O’Connor experiences a double transformative experience: a cognitive transformative experience, dementia, and a personally transformative experience, the loss of a loved one. My account suggests that his change of preference in his valuing of his life could arise from his neurodegeneration, or that it could arise as a result of having both of these transformative experiences. As we cannot know either way, my account suggests that we take Mr. O’Connor’s preferences seriously. That is, if Mr. O’Connor does get sick, has an advance directive which commits him to his prior religious beliefs, and is still maintaining he does not wish to be alive, it follows from my view that we ought to listen to Mr. O’Connor’s beliefs as a dementia patient and disregard his religious preference pre-dementia. According to my view, dementia patients can have positive or negative preference revelations, and both ought to be given moral weight by clinicians. Moreover, it seems that clinical practice has already moved in this way; my account is just shedding light onto why this may be.

I admit the possibility that there may be motivations which I have not adequately addressed in this article which remain legitimate reasons why an individual may desire to bind themselves to an advance directive. What measures can one take if, understanding the transformative nature of dementia, one still seeks to write an advance directive and bind oneself to it?

The following suggestions are tentative and have policy implications which cannot be fleshed out in sufficient detail in this article. Nevertheless, it is clear that my account favors stricter regulations on advance directive writing because the way that advance directives are currently written, and interpreted, is not helpful to clinicians who have to balance the considerations stipulated in an advance directive with the well-being of the patient they see before them. Therefore, I suggest that if one wants their advance directive to be implemented in clinical practice, one ought to be cautious both with respect to how one writes an advance directive and how one thinks their directive could be interpreted. With respect to writing, it is unadvisable for individuals to use websites which promote ‘easy’ advance directive writing, as the transformative experience framework has shown us, at the very least, that advance directive writing is not easy.[58] Consequently, it seems advisable for individuals in this position to seek guidance from both lawyers and clinicians, in order to increase the possibility that their directive be taken seriously in a clinical setting.

However, as the case of Mr. White illustrates, taking these steps with writing is not sufficient to ensure that one’s directive will be implemented. In order for this to be the case, one needs to ensure that the interpretation of one’s advance directive is relatively straightforward for medical professionals and loved ones. There is no easy means of achieving this end, but early discussions with medical professionals and loved ones could help. For instance, both parties

---

[58] My account gives us good reason to think that such websites should be banned.

need to know *how* binding you desire the advance directive to be, and, more importantly, *why* you want it to be binding. In Mr. White's case, it was not clear that, given the stage of his illness at the time the advance directive was drawn up, and the fact that he was only motivated to write it because his children desired him to, he could adequately answer either question. Likewise, it isn't clear from Mrs. Black's case how legally binding she desired her advance directive to be. Ambiguities such as these making interpretation difficult, if not impossible, for medical professionals. [59]

§5 Conclusion

This article has proposed a reason why one may want to have reduced confidence in writing advance directives. I began by presenting the received view put forward by Ronald Dworkin, which argues that advance directives ought to have significant moral weight when considering how to act in a dementia patient's best interests. I suggested that his account fails to consider an important variable which ought to hold moral weight when considering how to act in a patient's best interests. This variable is cognitive transformation. This article presents reasons why experiencing cognitive transformation can affect a person's beliefs, values, and preferences and why it is important to take these new preferences seriously. I suggest that this account can help to explain why it is that clinicians are reluctant to give moral weight to advance directives in medical decision-making. In so doing, my account aligns philosophical theory with current clinical practice.

---

[59] One tentative policy implication, then, is for clinicians and lawyers to actively discourage patients from using ambiguous phrases in the wording of their advance directives. In addition, it could be required that persons who write an advance directive answer both of these aforementioned questions in as much detail as possible.

Chapter 3 References

Albrecht. G. L., & Devlieger, P. J. (1999). The Disability Paradox: High Quality of Life Against All Odds, *Social Science and Medicine*, 48, 977-988.

Applebaum, P. S. (2007). Assessment of Patients' Competence to Consent to Treatment, *New England Journal of Medicine*, 357(18), 1834-1840.

Crisp, R. (2018). *The Oxford Handbook of Reasons and Normativity, ed. D. Star*, Oxford, UK: Oxford University Press.

de Boer, M. E., Hertogh, C., Dröes, R., and Jonker, C. (2010). Advance Directives in Dementia: Issues of Validity and Effectiveness. *International Psychogeriatrics*, 22(1), 201-208.

DeGrazia, D. (1999). Advance Directives, Dementia, and 'The Someone Else Problem'. *Bioethics*, 13(5), 373-391.

Dworkin, R.M. (1994). *Life's Dominion: Argument About Abortion and Euthanasia*. London: Harper Collins Publishers.

Fazel, S., Hope, T. & Jacoby, R. (1999). Dementia, Intelligence, and the Competence to Complete Advance Directives. *The Lancet*, 354, 9172.

Fehige, C. and Wessels, U. (2011). *Preferences*. Berlin, Germany: De Gruyter.

Fellows, L. (2015). Competency and Consent in Dementia, *Journal of the American Geriatrics Society*, 46(7), 922-926.

Fraassen, B. V. (1995). Belief and the Problem of Ulysses and the Sirens, *Philosophical Studies*, 77, 7-37.

Fried, T., Knight, S., McMahan, R., & Sudore, R. (2013). Advance Care Planning Beyond Advance Directives: Perspectives from Patients and Surrogates, *Journal of Pain and Symptom Management*, 46 (3), 355-365.

Groll, D. (2015) Medicine and Well-being. In (Eds.) G. Fletcher, *The Routledge Handbook of Philosophy of Well-Being*, London: Routledge.

Hawkins, J. (2014). Well-Being, Time, and Dementia, *Ethics,* 124 (3), 507-542.

Hope, T. (1994). Personal Identity and Psychiatric Illness. *Royal Institute of Philosophy Supplement*, 37, 131-143.

Hope, T., Slowther, A. and Eccles, J. (2005) Best Interests, Dementia and the Mental Capacity Act. *Journal of Medical Ethics*, 35(12).

Jaworska, A. (1999) Respecting the Margins of Agency: Alzheimer's Patients and The Capacity to Value. *Public Affairs*, 28 (2), 105–138.

Kitzinger, J. and Kitzinger, C. (2016) PPIW Report Increasing the Awareness and Uptake of Advance Directives. Retrieved June 17th, 2022, from http://ppiw.org.uk/files/2016/02/PPIW-Report-Increasing-the-awareness-and-uptake-of-Advance-Decisions.pdf

McMahan, J. (2003). *The Ethics of Killing: Problems at the Margins of Life.* New York: Oxford University Press.

Mitchell, W. (2018). *Somebody I Used to Know.* London: Bloomsbury Publishing.

Moye, J. and Marson, D. (2007). Assessment of Decision-Making Capacity in Older Adults: An Emerging Area of Practice and Research. *The Journals of Gerontology*, 62 (1), 3-11.

Nozick, R. (2001). *Anarchy, State, and Utopia*. Oxford: Blackwell Publishing.

Olson, E. T. (2019). *Personal identity*. Stanford Encyclopedia of Philosophy. Retrieved June 17, 2022, from https://plato.stanford.edu/entries/identity-personal/

Paul, L. A. (2016). *Transformative Experience*. Oxford: Oxford University Press.

Persad, G. (2019). Authority Without Identity: Defending Advance Directives via Posthumous Rights Over One's Body, *Journal of Medical Ethics;* 45: 249-256.

Sehgal, A., Galbraith, A., Chesney, M., Schoenfeld, P., Charles, G., & Lo, B. (1991). How Strictly Do Dialysis Patients Want Their Advance Directives Followed? *Journal of the American Medical Association*, 267 (1), 59-63.

Shiffrin, S. (2004). Autonomy, Beneficence, and the Permanently Demented. In (Eds.) J. Burley, *Ronald Dworkin and His Critics*, 195-217. New Jersey: Blackwell Publishing.

Sokolowski, M. (2018). *Dementia and the Advance Directive – Lessons from the Bedside. USA: Springer.*

The, A., Pasman, R., Onwuteaka-Philipsen, B., Ribbe, M., and Van der Wal, G. (2002). Withholding the Artificial Administration of Fluids and Food from Elderly Patients with Dementia: Ethnographic Study*, The British Medical Journal,* 325, 1326-1330.

Thomson, J. (1971). A Defense of Abortion. *Philosophy and Public Affairs*, 1 (1), 47-66.

Ubel, P., Schwarz, N., Loewenstein, G., and Smith, D. (2005) Misimagining the Unimaginable: The Disability Paradox and Health Care Decision Making. *Health Psychology*, 24 (4), S57-S62.

Vezzoni, C. (2005). *The Legal Status and Social Practice of Treatment Directives in the Netherlands.* Groningen: RUG.

Wells, C. (1977). *Dementia*. Pennsylvania, USA: Davis Company.

Chapter 4: Trusting at the End – Mosaic Trust and Dementia

The first three chapters of this thesis can be viewed as highlighting the negative personal and relational effects which arise from feeling stuck in time. Chapter 4 is instead an exploration of how one can potentially transcend these effects, even if they are stuck in time in the dual sense (both psychologically due to neurodegenerative illness and socially due to occupying a marginalized social position).

Whilst relations can hurt and hinder us, as this thesis has shown, they can also help us overcome some of the personal and relational harms associated with feeling stuck in time. Chapter 4 explores how, by showing that individuals with dementia may retain their ability for interpersonal trust even if they lose their capacity for self-trust. I do so by exploring the particular case of Iris Murdoch's dementia, showcased in the film *Iris*.

As indicated in Chapter 1, self-trust appears to be a precursor for one being able to successfully engage in MTT. Iris, in the moderate-late stages of dementia, appears to lack self-trust as her dementia has eroded her cognitive capacity for this particular kind of trust. Iris is thus stuck in time due to her dementia. Yet, Iris is able to transcend this feeling of being stuck through her relationship with John Bayley. Her lasting interpersonal trust in John enables her to mitigate some of the harms which usually arise from this phenomenology.

§1 Introduction

Trust in oneself and trust in others have often been conceived as requiring parallel conditions by trust theorists (Govier, 1993; McLeod & Sherwin, 2000). Both kinds of trust are theorized to comprise both cognitive and affective elements (Jones, 1996). The affective component of trust involves possessing a generally positive perspective, an ‘optimistic’ attitude that colors how I see myself and others (Jones, 1996). This optimistic attitude “restricts the interpretations we will consider as possibly applying to the words and actions of another” (Jones, 1996, p. 12). Imagine that I trust my friend to look after my apartment whilst away. In asking her to do so, I make several assumptions about her that I may not make if I asked a stranger to do the same job: I assume that she will treat my apartment with respect, that she will not steal from me, that she will contact me in any emergency and that she will refrain from setting the place alight in my absence. Affective dimensions thereby color my cognitive beliefs. I have faith in my friend because trusting her in the past has brought only rewards and no ill. Philosophers of self-trust have argued that the same can be said concerning trust in oneself (Govier, 1993). If I trust myself to show up to work on time, I have a generally positive outlook on my capacity to do that.

The cognitive component of trust involves beliefs in a person’s competency relative to the domain in question. In trusting my friend to look after my apartment, I believe her competent to look after an apartment. I would not trust her with this task if I did not believe she had the capacity to do it. Similarly, I may lack self-trust if someone asks me to look after their pet snake, having never cared for a snake before and therefore having doubts about my competency to do so. This competency requirement manifests itself differently in the various philosophical accounts of trust postulated in the philosophical literature. Still, the kernel of all accounts remains the same: the trusted individual must be deemed trust-responsive by the trustee (McGeer & Pettit, 2017). Trust-responsive can be domain-specific or generalized. If it is domain-specific, you assume the trusted individual will prove reliable for the assigned task. If it is generalized trust in the individual, you must believe your trust will prompt the trusted to act in specific ways (Jones, 2012; McGeer & Pettit, 2017).

This chapter aims to show that if we examine cases of severe cognitive impairment, such as the kind found in individuals with moderate-to-severe dementia, we begin to see that self-trust does not neatly mirror interpersonal trust: the two can, and do, come apart. I will argue that although self-trust requires a cognitive component, interpersonal trust does not necessarily require this. This shows that we need a philosophical account of trust to explain why individuals with cognitive impairments may lose self-trust but retain interpersonal trust. Individuals in the moderate-to-severe stage of dementia do not appear to reason based on beliefs in another's trust responsiveness: they do not necessarily seem aware of holding these beliefs and, even if they do, often lack the tools to articulate such beliefs. Therefore, on standard philosophical accounts of trust, these individuals do not meet the competency threshold to qualify as being in relations characterized by interpersonal trust. However, individuals who may not meet the cognitive requirement for trust could still be in trusting relations with others in their social milieu. Such examples show that we need a philosophical account to explain why individuals with cognitive impairments may lose self-trust but retain interpersonal trust. My suggestion is that original philosophical accounts of trust as an "affective atmosphere" more than a cognitively complex phenomenon, such as Baier's (Baier, 1986) and Bok's (Bok, 1978), can provide us with such an account. Such accounts, however, do not elucidate what this affective atmosphere may be constituted by. I propose that we conceive of the affective atmosphere between two people as being a mosaic, comprised of the individual’s behavioral dispositions, beliefs, and emotional states. This mosaic is constitutive of the affective atmosphere which I propose as the sufficient condition for interpersonal trust. It can also provide the means by which medical professionals could judge an individual with severe cognitive impairment to be in interpersonal relations of trust with others.

The chapter will be structured as follows. I begin by articulating different accounts of self-trust and interpersonal trust, which will allow us to evaluate them in the case study of dementia. I then flesh out what dementia is and how this condition affects one’s ability to trust oneself and others by employing the film *Iris* (Eyre, 2001) based on Iris Murdoch’s dementia, retold by her husband, John Bayley. The film portrays how dementia warps Murdoch’s capacity to engage in her usual activities. Nevertheless, it also highlights that, even amid severe dementia, she was able to trust John, poignantly even as she loses trust in herself in the moderate stage of the disease. Such a case problematizes both the assumption that self-trust neatly maps onto interpersonal trust and the idea that individuals need to hold cognitively complex beliefs, such as I believe X is competent in domain Y for reason Z. I show that the behaviors Iris displays toward John indicate a generalized trust in him, which cannot be reduced to beliefs regarding the trust-responsiveness of John, because it is not clear that Iris, and others in Iris’s position, are capable of holding such complex beliefs. I argue that these cases suggest that trust can be best captured by concepts such as an affective atmosphere of trust and a generalized sense of goodwill (Bok, 1978; Baier, 1986).

However, the question remains how to understand what an affective atmosphere of goodwill between two people amounts to. To address this question, I suggest we utilize Fins’ account of the mosaic medical decision-making (Fins, 2017). Fins’ account of mosaic medical decision-making captures the idea that all aspects of the person's behavior have something to contribute to the decision-making process regarding individuals who are in minimally conscious states (Fins, 2017). Fins suggests that we should bring "different shards of information together to construct a coherent picture, much the way small stones coalesce into a beautiful mosaic" for such individuals (Fins, 2017, p. 168). I argue that trust can work similarly in cases of severe cognitive impairment. Therefore, I propose that the affective atmosphere between two people is constituted by a mosaic of emotional states and behavioral dispositions and that this is sufficient for interpersonal trust. This mosaic additionally provides one means to judge that the person with severe cognitive impairment trusts another.

§2 Accounts of Interpersonal Trust and Self-Trust

Interpersonal Trust

One of the foundational thinkers regarding interpersonal trust is Annette Baier (Baier, 1986). She proposes a *goodwill* account of self-trust (cf. Jones, 2012) because, unlike many of her contemporaries, she emphasized the importance of the affective element of interpersonal trust such as the idea that trusting essentially involves entrusting others with domains we care about. Baier begins by acknowledging that there had been a myopic approach to trust in the philosophical literature before her account, due to the philosophical obsession with individualism (Baier, 1986). She points out that we do not live individualistic lives, but rather inhabit networks and a system of connections with others, many of which take place in an "atmosphere of trust". Much like breathing, we only notice this atmosphere when we are struggling to experience it or when trusting relations become scarce and polluted (Bok, 1978; Baier, 1986). Baier identifies the relationship between infants and their parents as one of trust: "[s]ome degree of innate, if selective, trust seems a necessary element in any surviving creature whose first nourishment (if it is not exposed) comes from another" (Baier, 1986, p. 242). According to Baier, there can be conscious forms of trust, such as those between competent adults, and unconscious forms of trust, such as the kind shown between infants and children.

Her account suggests that trust depends on goodwill towards one from another, insofar as "[w]hen we are trusted, we are relied upon to realize *what* it is for whose care we have some discretionary responsibility" (Baier, 1986, p. 236). She gives the example of a babysitter who decides to transform the nursery of the child they are looking after. Baier points out that she has overstepped her caring responsibilities, as she was entrusted with the act of looking after the child, not decorating the nursery. Baier's account of interpersonal trust emphasizes that "Trust… is reliance on others' competence and willingness to look after, rather than harm, things one cares about which are entrusted to their care (Baier, 1986, p. 259). It has, therefore, an element of competency, but this appears to be minimal compared to the philosophical accounts of trust that arose subsequently.

In response to Baier, Karen Jones (2012) rejects a goodwill account of trust. She adopts a three-placed notion of trustworthiness on which "B is trustworthy with respect to A in domain of interaction D if and only if she is competent with respect to that domain, and she would take the fact that A is counting on her to be a compelling reason for acting as counted on" (Jones, 2012, p. 61). On such an account, for me to act as trusting you in a particular domain, it would be necessary for you to be competent to understand your intentions as well as my own. Jones thereby shifts the lens from the affective components of trust to the cognitive, emphasizing the importance of competency in the domain in question.

Most recent theories of philosophical trust outline an empowering theory of trust intended to capture the importance of both cognitive and affective dimensions of trust (McGeer & Pettit, 2017). The empowering theory of trust suggests that trust is a function of two variables: how trustworthy you are and how far you are moved by the presence of those who manifestly rely on you (McGeer & Pettit, 2017, p. 21). Such an account, whilst it involves a discussion of the affective importance of trusting others, namely that trusting others can promote others to be more trustworthy as individuals strive to meet the expectations of the relationship, places most emphasis on the cognitive components of trust. They highlight the central importance of the trusting individual having a belief in the trusted person's capacity in the relative domain (McGeer & Pettit, 2017). Similarly to Jones' account, trusting requires complex cognitive beliefs in the trustworthiness of the trusted individual.

As fleshed out, it is clear that current philosophical accounts of trust typically conceive of trust as requiring the trustee to recognize and judge the competency of the trusted. This occurs with all the accounts presented, including Baier's, albeit to a lesser degree than the other two accounts. This competency requirement instantly puts individuals who suffer from severe cognitive impairment in a position where they may not be able to meet the requirement necessary for interpersonal trust because individuals in this position are rarely in a position to judge another's competency in this way. And even if they retain such an ability, they may lack the ability to communicate their judgements.

Self-Trust

Accounts of self-trust typically begin with the assumption that self-trust is akin to interpersonal trust. Such accounts are thereby modelled on accounts of interpersonal trust. Self-trust theorists suggest that similarly to interpersonal trust, trust in oneself depends on having a general disposition to view oneself favorably. This does not mean that in all circumstances, one has to view oneself in the most favorable light or that one must have confidence in one's abilities that are not yet developed. It does not mean, for instance, that I lack general self-trust simply because I lack trust in my ability to code computers, for which I have no training. Rather, self-trust has been commonly understood in more general terms as "the representation of the self as able to cope with the world and itself" (Brothers, 1982). This process also involves cognitive and affective elements including a positive view of oneself and belief in one's competency (Govier, 1993). Self-trust is characterized as involving four core elements: (i) one having positive beliefs about one's motivations and competence, (ii) seeing oneself as a person of integrity, (iii) being willing to rely on oneself, accepting relative risks and vulnerabilities in doing so, and (iv) having a general disposition to see oneself in an optimistic light (Govier, 1993, pp. 105-106). Govier argues that when these attitudes exist in relation to oneself, she can be said to have self-trust.

On McLeod & Sherwin's account of self-trust, self-trust "is an attitude that is shaped by beliefs and/or feelings about the trustworthiness of the trusted", namely oneself (McLeod & Sherwin, 2000, p. 265). On their account, self-trust is required for an agent to count as fully autonomous. McLeod & Sherwin identify three types of self-trust: Type I, Type II, and Type III. Type I self-trust "involves having good decision-making skills and also being situated to choose well" (McLeod & Sherwin, 2000, p. 263). Type II self-trust is one's ability to act on the decisions one makes (McLeod & Sherwin, 2000, p. 264). Type III self-trust requires that the person must trust the judgements that underlie her decisions (McLeod & Sherwin, 2000, p. 264).

Self-trust, as conceptualized, is threatened by dementia, as will be shown below. However, I will argue that interpersonal trust is not as threatened and can survive, even for people with severe cognitive impairment. The next section will show how self-trust and interpersonal trust can come apart in cases of individuals with compromised cognitive functioning and then explore what this shows us about interpersonal trust.

§3 Dementia, Trusting Oneself, & Trusting Others

Dementia has often been viewed through "a lens of disabling symptoms, deficit and risk" (Clark, Phillipson, & Ward, 2021), owing to the fact that psychiatrists have typically conceptualized dementia as an illness of deficits. Dementia may affect the following: orientation, affectivity, relationality, memory, autonomy, intellectual function, creativity, and judgement. Individuals with dementia do not know which area they will retain control over and which cognitive site will degenerate first. Symptomology depends on the variant of dementia facing the individual. The most common forms of dementia are Alzheimer's disease, vascular dementia, dementia with Lewy bodies, and frontotemporal dementia (Gale, Acar, & Daffner, 2018). These dementias can be further split into two groups: neurodegenerative and non-neurodegenerative. Alzheimer's disease is a common form of the former, and vascular dementia is a common form of the latter.

Neurodegenerative dementia has been medically conceptualized as being comprised of three stages, which vary with the kind of dementia involved. These three stages are: early, moderate, and severe. Early-stage dementia patients, particularly those with Alzheimer's disease, typically struggle with a loss or change in short-term memory and intellectual difficulties such as "patients may show subtle forgetfulness, or occasionally repeat stories and can exhibit irritability, apathy, or low mood" (Gale et al., 2018). For example, individuals may struggle to remember their timetable or more easily forget their words. In the moderate stage, dementia patients become less preoccupied with their previous ambitions and appear to care less about social responsibilities that once mattered greatly to them (Wells, 1977). Finally, the last stage of the disease is when full-time care for the individual is required, as the individual can suffer from a lack of orientation and ability to communicate, and a decline in motor, spatial, and sensory functioning (Clark et al., 2021). Typically, patients can exhibit 'wandering' behaviours in this stage of the disease, such as escaping from their homes. An exploration of how neurodegenerative conditions such as dementia affect one's capacity for self-trust and interpersonal trust is still underexplored in the philosophical literature.

I now turn to a discussion of Iris Murdoch's life and her dementia, depicted in the film *Iris*. I argue that the presentation of Iris's dementia is suggestive that self-trust and interpersonal are not parallel, as assumed in the philosophical literature, because Iris appears to lose self-trust but retains interpersonal trust in her husband, John Bayley.

> INTERVIEWER
> When did you know you wanted to write?
>
> MURDOCH
> I knew very early on that I wanted to be a writer. I mean, when I was a child I knew that. Obviously, the war disturbed all one's feelings of the future very profoundly. When I finished my undergraduate career I was immediately conscripted because everyone was. Under ordinary circumstances, I would very much have wanted to stay on at Oxford, study for a Ph.D., and try to become a don. I was very anxious to go on learning. But one had to sacrifice one's wishes to the war.

(Meyers, 2013)

It is important to remember who Iris was before evaluating how Iris's dementia affected her self-trust and interpersonal trust. Iris was born in Dublin, Ireland, on July 15th, 1919. Her family moved from Dublin to London when Iris was young (Blum, 2022). Iris's literary life began at only thirteen years of age, in 1932, when she was awarded one of the first scholarships to the liberal Badminton School in Bristol (Martin & Rowe, 2010). At this young age, Iris was already writing on philosophical topics such as "How I Would Govern the Country", arguing against totalitarianism (Martin & Rowe, 2010). The years 1938-1946 were formative for Iris's philosophical interests, as Iris met Phillipa Foot, Mary Midgley, and Frank Thomson during her undergraduate studies at Oxford (Martin & Rowe, 2010). During WWII, Iris made herself of service to England, taking up roles as a civil servant in London in 1942, then working for the United Nations Relief and Rehabilitation administration in 1944. After the war, Iris resumed her previous studies, taking up a research studentship in philosophy at Newnham College, Cambridge (Martin & Rowe, 2010). The following year Iris took up a position as a fellow and tutor at St. Anne's College, Oxford, where she began publishing essays and making BBC broadcasts (Martin & Rowe, 2010).

Iris's first book in philosophy "Sartre: Romantic Rationalist" was published in 1953 (Blum, 2022). Her philosophical interests ranged from phenomenology, analytic philosophy, contemporary existentialism, Hindu and Buddhist philosophy, to moral philosophy. She is also celebrated for her unique contribution to Platonic moral realism and her distinctive ethics. Iris's literary career was as pronounced as her academic career, as she became known as one of the most prominent novelists of the 20th century (Martin & Rowe, 2010). Iris published twenty-six novels and six plays over her lifetime (Meyers, 2013). Her first novel, published in 1954, entitled *Under the Net,* won the runner-up prize at the Cheltenham literature festival (Martin &

Rowe, 2010). Her novel *The Sea, The Sea* went on to receive the Booker Prize. Iris's ability to mix philosophy and fiction was admirable. Her ability to speak to audiences comprised of individuals from all backgrounds and all walks of life resulted in many of her novels being classified as bestsellers (Martin & Rowe, 2010).

Moreover, Murdoch's career as both a philosopher and a novelist resulted in profound love and appreciation of Iris in England, so much so that when she passed from dementia on the 8th of February 1999, at Vale House in Oxford, the BBC news prioritized reporting her death over the death of King Hussein of Jordan (Martin & Rowe, 2010). That was her importance then, but the afterlife of her works lives on in England and other parts of the world, growing in significance past her death. Iris's novels continue to be viewed as works full of "sensuous pleasures, fantastic invention, high intelligence, and moral dignity" (Meyers, 2013).

We are now able to explore the moments leading up to Iris's death, the impact dementia had on Iris's capacity for self-trust, and her capacity for interpersonal trust.

Self-Trust

> It's only the postman
> It's only the postman
> It's only the postman
>
> John, which side do I go?
>
> (Eyre, 2001)

Dementia clearly affects one's capacity to navigate one's relationship with oneself, with others, and with one's environment. It affects competency by causing individuals to gradually lose their ability to do things they once could do with ease. This erodes self-trust, given that self-trust is a perception of one's ability to cope with the world. Dementia affects one's ability to manage in the ways one ordinarily would. Consider the early stages of Iris's dementia, in which, as John notes, she began struggling with her speaking and writing. As a philosopher and author, she relied on her capacity to write well and speak fluently. Upon first suffering from symptoms of dementia, the first aspect of competence that Iris grapples with losing is the capacity that she has relied on for her career. Writing her last novel, Iris reports struggling to find her words; she struggles to convey her ideas in the coherent way she is used to. This, in turn, severely compromises Iris's ability to have self-trust, as she no longer feels able to cope with the activity she loved doing the most and had the most competence in being able to complete.

Individuals living with dementia are also not always 'trustworthy' in the sense that matters for self-trust (McLeod & Sherwin, 2000), particularly in moderate-severe stages of the disease. In the moderate stage of dementia, Iris begins to repeat phrases, forgetting she had just said them and does not know how to do things such as walk through the doorway, which is why she asks John, "Which side do I go?" (Eyre, 2001). She relies on John to be able to tell her she is repeating herself and to help her navigate her environment as dementia begins to affect her speech, movement, and perception. In these ways, Iris's illness causes her to lose self-trust in completing everyday tasks alone because she comes to lack belief in her competency to do these tasks. She also begins to lose the affective dimension of self-trust: the positive conception of herself and her abilities. She realizes she is frustrating John by repeating herself, but she cannot stop because she cannot remember *why* she is frustrating him. During the moderate stage, she tries to apologize, knowing that she is doing something wrong but unsure of precisely what, making the sounds for the beginning of the word "Sor..". She cannot finish her sentence, to which John responds, "Please don't, Iris" (Eyre 2001).

Individuals with dementia are sometimes aware of their lack of self-trust and at other times, lack such lucidity. This happens most starkly in the severe stage of dementia for Iris. Even when in the severe stages of dementia, Iris has moments of lucidity where she remembers that she loved writing. She says to John, "I wrote.." unable to finish the sentence, to which John replies, "Such things you wrote" (Eyre, 2001). This causes John to speculate to Iris's doctor that her map of the world cannot be wholly eaten. John notes:

> You have shown me a map of Iris's brain world. Empty. You tell me that all its hidden mysteries and all the unknown life in there has gone…How then can she say things of such terrible lucidity? Which bit of the empty jungle does that come from? Is it not remotely possible that what Iris says is of some consequence? Perhaps we ought to learn her language before the lights go out…
>
> (Eyre, 2001)

Due to how dementia disrupts Iris's capacity for self-trust, John is shocked when Iris remembers things and expresses knowledge in certain areas. This is poignantly demonstrated in the film when John has become unable to care for Iris and has to have her put into care. Iris sits on the staircase, refusing to leave John, which indicates on some level that she still wants to stay with John and cares about being in her own house. The taxi driver, who was a member of the caring staff at her new care home, trying to encourage her to move from the stairway and come with him, asks, “What’s the name then?” to which Iris replies quite promptly, given her deteriorating mental state, “Iris” (Eyre, 2001). Thus, even though most self-trust is eaten away in severe dementia patients, the localization of self-trust is an essential element: they may lack self-trust in their ability to walk through doorways but remember whom they love and their name when asked. Other individuals living with dementia may experience the reverse: they may retain their ability to do simple tasks such as walk into spaces and draw whilst being unable to remember their names or their loved ones’ names. Regardless, self-trust does appear eroded in neurodegenerative conditions and conditions which cause one to no longer feel confident in one’s ability to do the task one is required to do. This is because such conditions affect how capable one is in dealing with the world around one, which is a core component of self-trust on all accounts (Govier, 1993; McLeod & Sherwin, 2000).

Interpersonal Trust

> Iris: Help, help, help!!
> John: I am trying to help you (bursts out crying)
> Iris: So…so… (attempting to say sorry)
> John: Please don’t, Iris
>
> (Eyre, 2001)

The interesting question is whether dementia also erodes Iris’s capacity to trust John in the same way it affects her ability to trust herself. Watching *Iris,* it is clear that a similar erosion doesn’t occur in all instances of dementia. John’s trust in Iris to continue writing and speaking helps her reconcile herself to the deterioration of her capacities. Iris appears to trust John more than ever in her deteriorating state; she continues to demonstrate through her words trust in his ability to help her, calm her, and love her.

This is not to say that these trusting relations are not punctuated with difficulty and pain for the individual trusted by the trustee with a condition such as dementia. Those living with dementia may lose their ability to control their behavior in ways viewed as challenging by their primary caregivers. In Iris's case, this showed up most prominently in the moderate-to-severe stage of the disease, where she could not regulate her ability to go to the toilet and her ability to stop herself from 'wandering'. These behaviors considerably strain John, who begins initially by responding to these events in a compassionate way. When Iris relieves herself in the living room, rather than the bathroom, John replies, "You are becoming quite the naughty cat Iris", which she responds to with a smile (Eyre, 2001). Later on, though, after undertaking continuous caring responsibilities during day and night, John ends up screaming at Iris, becoming frustrated with her humming in the middle of the night, which is keeping him from resting:

> I hate you. Every bloody inch of you.
> All of your friends have finished with you.
> I've got you now, and I don't want you.
>
> (Eyre, 2001)

Upon watching John's anger build, Iris leans over to cuddle him, trying to calm him down. This presentation of Iris's reaction to John indicates that philosophical accounts of trust that rely on competency criteria are not adequate to capture the relations of trust genuinely experienced by those with diminishing capacities, such as those with neurodegenerative conditions. If Iris lacked trust for John, it would have been demonstrated in this scenario. If she didn't know who John was or how he cared for her, she would not have attempted to cuddle him; she would have been afraid and would have tried to leave the room or escape the situation. Yet, even when in the severe stages of the disease, Iris does not do this. Iris trusts John even in his angry state, knowing that he is trying his best to help her, not hurt her. Iris demonstrates similar amounts of trust and love for her friends, screaming at her friend's funeral "No! No! No!" and bowing to her other friend, whose name she cannot remember (Eyre, 2001).

There are two responses that current accounts may give to explain Iris's behavior. The first is to consider whether Iris's behavior demonstrates reliance rather than trust. Jones, for instance, states:

> I can rely on someone to behave in a certain kind of way because I have evidence that it is likely that she will behave in that way out of, say, habit, fear, vanity, or stupidity.
>
> (Jones, 1996, p. 15).

However, if Iris relied on rather than trusted John, we would expect her to react negatively to his aggressive outburst. Iris recognizes that, even in an agitated state, John has a particular disposition towards her, a disposition of goodwill, which she brings out by cuddling him, and at times apologizing to him. And John realizes that, even in a state of severe dementia, Iris loves him. This sentiment regarding the disposition of goodwill that individuals with cognitive impairment can have is expressed powerfully by Eva Kittay, who reminds us: “what it is to be human is not a bundle of capacities. It’s a way that you are in the world, a way you are with another” (Kittay, 2010, pp. 407-408).

Those who subscribe to other accounts may also respond that Iris still believes in John’s trust-responsiveness but cannot articulate this belief meaningfully. However, it is not clear, given Iris’s presentation during the severe stages of her dementia, that she does have the capacity to retain beliefs of the kind necessary for contemporary accounts of trust such that she is sure John is trust-responsive (Jones, 2012; McGeer & Pettit, 2017). Iris cannot remember what happened several minutes before, so this explanation does not seem to capture what is happening in this case.

Thus, I propose instead a return to Baier’s affective account, which suggests that there can be an affective atmosphere of trust between two individuals, which can be conscious or unconscious. Recall that the affective component of trust according to Baier is that one has a disposition of goodwill towards the one who is being trusted. Iris demonstrates this clearly in her response to John’s outburst of anger, by responding to his anger with compassion. Whilst Baier emphasized that trust is often characterized as believing in another’s competency to do something, I suggest her account works equally well if we remove this requirement for individuals with compromised cognition. Iris, after all, may have lacked the competency required to know that John would not hurt her despite his anger at her. Yet, Iris appears to recognize John’s goodwill towards her even if she can’t hold or retain specific beliefs about his trustworthiness.

Baier herself seemed to acknowledge that such cases of this kind of interpersonal trust are plausible, as she details the ways in which infants demonstrate trust in their caregivers, even though they lack the tools to communicate and are otherwise socially vulnerable. We know an infant trusts another in how they respond to the other in question. As Baier notes, “[i]nfants too can make suspicious, futile, self-protective moves against the powerful adults in their world” (Baier, 1986, p. 241). However, the question remains as to what an affective atmosphere of goodwill between two people looks like. I subsequently argue that this affective atmosphere is compromised by a mosaic of the individual's behavioral dispositions, emotional states, actions, and loving orientations.

§4 Mosaic Medical Decision-Making as a Framework for a Philosophical Account of Mosaic Trust

First, some background is required regarding the framework of mosaic medical decision-making that I will use as a model. Mosaic medical decision-making has been proposed for patients re-emerging from states of minimal consciousness (Fins, 2017). Goering suggests this should apply to patients who may never recover their cognitive capacities (Goering, 2020, p. 79). Mosaic medical decision-making is defined as an approach

> which incorporates the views of the surrogate decision-maker and the patient’s prior wishes and current articulation, as well as the input of a medical professional and a patient advocate…this group brings different shards of information together to construct a coherent picture, much the way small stones coalesce into a beautiful mosaic.
>
> (Fins, 2017, pp. 164, 168)

Fins' project is, in many senses, the reverse project of my own. Fins aims to provide a means to evaluate how patients can start to re-communicate when they are experiencing re-emergent agency after being in a minimally conscious state (MCS), whereas my focus is on agents who are in the process of losing their agency. Yet, we share the same goal in that we both wish to avoid the silencing of people with cognitive impairments who have diminished agency:

> These individuals should not be silenced again because of scholarly inattention to the pressing challenge of how to allow them to participate in decisions and discussions that bear on their welfare and happiness. Along with improved prosthetics, these individuals are also entitled to a suitable and workable ethical framework that would allow them the opportunity to demonstrate their fullest will consistent with prudence.
>
> (Fins, 2017, p. 165)

Similarly to MCS patients, I suggest that individuals with dementia are often left out of participation in decision-making that affects them simply due to ignorance about how best to include them in discussions concerning their welfare.

I propose here that Fins' insights apply to individuals living with dementia. His approach aims to emphasize three factors: inclusion, temporality, and overcoming fragmentation. Fins argues that the goal of medical decision-making ought to be valuing all perspectives; we have a moral imperative to include the voices of those who lack cognitive capacities. This act of inclusion takes place across time. Fins notes that the process of inclusion is ‘constructive’ in that caregivers, medical staff, the individual themselves, and family members must continue to build together to act in the individual's best interests. Mosaic approaches emphasize that no party “has full normative authority or substantive knowledge to decide alone” (Fins, 2017). Finally, mosaic medical decision-making is designed to overcome fragmentation, which can result from suffering from cognitive compromise of any kind. According to such a framework, individuals who are recovering and regaining cognitive capacities deserve to be integrated back into society again (Fins, 2017). No one is allowed to disvalue or dismiss the current preferences of the patient instantly, even if they struggle to understand them. The patient’s insistence on a choice is indicative that it holds value for them and ought to be paid attention to. It is also essential to pay attention to what the MCS patients say and what they do: their actions.

Fins' suggestion that individuals who are regaining conscious awareness deserve to be included, allowed to construct and have a voice in their interests, and be integrated into society is powerful. I argue that we should also care to give that treatment with dignity and respect to individuals who are losing capacities. In particular, I suggest that Fins’ account can ground a new philosophical understanding of interpersonal trust: mosaic trust. The mosaic approach to trust helps to explain what constitutes an affective atmosphere of goodwill between persons. It can also help medical professionals and family members of individuals living with severe cognitive impairment know if interpersonal trust is present. The mosaic theory of trust assumes that individuals with even minimal consciousness who may not know precisely how competent another is in a particular domain have an affectively rich internal life and may retain their capacity for interpersonal trust despite their cognitive impairment. As shown by Fins, individuals in MCS still attempt to communicate and still have ideas; these are just seldom taken into account because the patient with MCS may lack the tools to express them. Lacking the tools necessary to share one's perspectives and preferences does not indicate that one doesn't have any. I argue that the same is true with respect to interpersonal trust. Even if expression and communication are impaired due to cognitive impairment, that alone cannot be taken to suggest that there is no affective atmosphere of trust present. As demonstrated by the case of Iris, those with severe cognitive impairment can still be in a position to judge the goodwill of others towards them if we are willing to construct a mosaic of their behavioral dispositions and emotional states.

The mosaic theory of trust is significant because it answers how we are to understand the affective atmosphere which ought to be deemed sufficient for interpersonal trust. In everyday instances of trusting, it isn't clear that individuals always hold cognitively complex beliefs regarding those they trust. It seems equally plausible to suggest that the trusting can be reduced to the mosaic of the individual's behavioral dispositions, emotional states, and loving orientations towards others in their life. In this way, the mosaic of these dispositions, states and orientations is constitutive in cases of interpersonal trust.

The mosaic theory of trust also gives us the tools to begin judging if the individual with cognitive impairment believes the individual in question to have an attitude of goodwill towards them. Applied to the case of individuals with dementia, the mosaic account demands that we look at how the individual is behaving, what they are saying, and what they attempt to communicate through actions or words as to whether they trust the individual in question. On such a theory, we ought not to rule out the possibility that Iris has interpersonal trust in John due to her cognitive impairment until she has demonstrated through a combination of her actions, words, and behaviors that this is the case. If Iris shows affection towards John, if she appears to realize that John has goodwill towards her even if she cannot articulate it, we should be inclined to believe she retains trust in John.

This account of judging interpersonal trust between an individual with capacity and an individual with severe cognitive impairment has several benefits. One of the primary benefits is that it helps to explain the trust between the mentally ill and the mentally well; those losing their capacities and those in control of the same capacities; the dying who, though they lack cognitive lucidity, still turn to those they love in acts of trust. It does so without need or reference to having the mental awareness of one's beliefs of another's trustworthiness. We know these trusting relations exist, yet we cannot make sense of these relationships with the strict demands placed on "trustworthiness" in the current philosophical literature.

§5 Reflections on the Mosaic Account of Trust

Although self-trust is taken to be akin to interpersonal trust, the two can and do come apart in instances of individuals with severely compromised cognitive capacity. As argued in this chapter, it is clear that self-trust relies on a positive affective view of one's competence concerning some task or domain, namely that one's capacities regarding that task or domain are viewed in a positive light. To have a positive affective conception of oneself, one needs to feel capable of engaging in the activity in question. Individuals with moderate-to-severe dementia lack this capability and, at lucid moments, are aware of this. Such individuals lose self-trust because they lose the belief in their capability. Thus, while self-trust appears to have a cognitive element, it is sufficient for interpersonal trust that there is an interpersonal affective atmosphere of goodwill which can obtain even if individuals lack the complex beliefs ordinarily required for trust (such as that the person trusted is trust-responsive). This account suggests that essential human capacities, for affection and loving, and for desiring to communicate, may be present even when self-trust is severely compromised or lost altogether.

The mosaic account suggests that the presence of interpersonal trust may help repair some of the damage done by lost self-trust, for instance when a person with dementia is included in conversation rather than excluded and when one attempts to speak their language rather than talk over them. Maintaining interpersonal trust helps to keep the person connected, by seeking to understand rather than dominate, and by identifying that those with profound cognitive disabilities have something to offer, something to share with us and the world, that we may miss if we are not in relation with them (Brison, 2021). The mosaic account of trust suggests that, by focusing exclusively on one's capacity for cognitive reasoning and trust-responsiveness, we will ignore a whole range of individuals in relationships characterized by interpersonal trust by recognizing competency as a central idea.

Dementia affects both episodic memories, the kinds we have as individuals regarding our pasts, and semantic memories, the memories we have of know-how about the world (Irish, 2016). Both kinds of memory are central to "mental time travel"; the faculty used to re-live events and project forwards in time (Suddendorf & Corballis, 2007; Michaelian, 2016). One upshot of the mosaic account is that mental time travel may be a capacity which can be scaffolded through our affective relations with others. The scaffolded mind refers to the idea that our minds are environmentally supported; specific environments can increase and amplify a person's cognitive capacities, whereas others appear to diminish them (Sterelny, 2010; Colombetti & Krueger 2015). Take, for instance, the comfort one feels in sitting down to write in their favourite café (Sutton, 2018). My proposed account encourages us to view intimate relationships to work in a similarly supportive way. John's trust in Iris scaffolds who she believes she is. And even in late-stage dementia, Iris was able to scaffold John by responding to him in kindness when angry. Philosophers who defend accounts of interpersonal scaffolding of affectivity emphasize that trust can be construed as “the expectation that others will have a certain modulatory impact on our affective life” (Colombetti & Krueger, 2015, p. 1166). The mosaic account of trust suggests that individuals who cannot articulate the effects that others have on them may still feel and be affected by these effects. Thus, this account favours returning to the affective heart of trust and focusing less on the meta-cognitive abilities required for trusting.

This chapter has argued that we ought to ground philosophical accounts of trust in Murdoch’s suggestion that loving is an orientation, not simply a state of mind. Individuals with cognitive impairments may lack the state of mind necessary for recognizing another’s trust-responsiveness, but they do not necessarily lack a direction for love, as *Iris* has shown us. I argued that individuals with cognitive impairments can be in trusting relationships with others, and we need a philosophical account that makes sense of how this is possible. To include individuals with compromised cognitive capacities as being able to trust others, I proposed to adjust Baier’s original conception of goodwill to remove the competency requirement. Recall that Baier suggests that to trust another is not only to have a sense that they are motivated by goodwill in a particular domain, but also to have a belief in their competency in that domain. My suggestion is that the second condition is not necessary for individuals with severe cognitive impairments. To explicate trusting relations, I proposed a theory of mosaic trust, whereby we assess what the trustee is saying, doing, and behaving to indicate their trust in the trusted. John’s relationship with Iris has revealed that there is more than one sense in which trusting someone who is losing trust in themselves can give them the love needed to move through life, with whatever capacities they have.

Chapter 4 References

Baier, A. (1986). Trust and Antitrust. *Ethics* 96 (2): 231–60.

Blum, L. (2022). *Iris Murdoch.* Stanford Encyclopedia of Philosophy. Retrieved 17th 2022, from https://plato.stanford.edu/entries/murdoch/

Bok, S. (1978). *Lying: Moral Choice in Public and Private Life*. New York: Pantheon Books.

Brison, S. J. (2021). Valuing the Lives of People with Profound Intellectual Disabilities. *Philosophical Topics*, 49 (1): 99–122.

Brothers, D. (1982). Trust Disturbances in Rape and Incest Victims (PhD Thesis). New York: Yeshiva University.

Clark, A., L. Phillipson, & R. Ward. (2021). *Dementia and Place*. Bristol: Bristol University Press.

Colombetti, G., and J. Krueger. (2015). Scaffoldings of the Affective Mind. *Philosophical Psychology* 28 (8): 1157–76.

Eyre, R. (Director). (2001). *Iris [film]*. BBC Films: Fox Iris Productions.

Fins, J. J. (2017). Mosaic Decision-making and Reemergent Agency After Severe Brain Injury. *Cambridge Quarterly of Healthcare Ethics*, 27 (1): 163–74.

Gale, S. A., D. Acar, and K. R. Daffner. (2018). Dementia. *American Journal of Medicine*, 1161–69.

Goering, S. (2020). Why Should Adamancy of an Uninformed View Give Moral Weight? *The American Journal of Bioethics* 20 (8): 78–79.

Govier, T. (1993). Self-Trust, Autonomy, and Self Esteem. *Hypatia,* 8 (1): 99–120.

Irish, M. (2016). Semantic Memory as Essential for Future-Orientated Mental Time Travel. In (Eds.) K. Michaelian, S. B. Klein, K. K. Szpunar, *Seeing the Future.* Oxford: Oxford University Press.

Jones, K. (1996). Trust as an Affective Attitude. *Ethics* 107 (1): 4–25.

Jones, K. (2012). Trustworthiness. *Ethics* 123 (1): 61–85.

Kittay, E. (2010). “The Personal is Political is Philosophical: A Philosopher and Mother of a Cognitively Disabled Person Sends Notes from the Battlefield.” In (Eds.) E. F. Kittay & L. Carlson, *Cognitive Disability and Its Challenge to Moral Philosophy,* 393-413. Malden, MA: Wiley-Blackwell.

Martin, P., and A. Rowe. (2010). *Iris Murdoch: A Literary Life*. London: Palgrave Macmillan.

McGeer, V., and P. Pettit. (2017). The Empowering Theory of Trust. In (Eds.) P. Faulkner & T. Simpson, *The Philosophy of Trust*. Oxford: Oxford University Press.

McLeod, C, and S. Sherwin. (2000). Relational Autonomy, Self-Trust, and Health Care for Patients Who Are Oppressed. In (Eds.) N. Stoljar and C. Mackenzie, *Relational Autonomy*, pp. 259–79, Oxford: Oxford University Press.

Meyers, J. (2013). *Remembering Iris Murdoch: Letters and Interviews*. New York: Palgrave Macmillan.

Michaelian, K. (2016). *Mental Time Travel*. Cambridge, MA: MIT Press.

Murdoch, Iris. (2003). *Metaphysics as a Guide to Morals*. Vintage Classics. London: Vintage Books.

Sterelny, K. (2010). Minds: Extended or Scaffolded? *Phenomenology and Cognitive Sciences* 9: 465–81.

Suddendorf, T., and M. Corballis. (2007). The Evolution of Foresight: What Is Mental Time Travel and Is It Unique to Humans? *Behavioural and Brain Sciences* 30 (3): 299–313.

Sutton, J. (2018). Shared Remembering and Distributed Affect. In (Eds.) K. Michaelian, D. Debus, and D. Perrin, “*New Directions in the Philosophy of Memory”.* Oxford: Routledge.

Wells, C. (1977). *Dementia*. USA: Davis Company.

A Comprehensive Scholarly Discussion of the Findings

This thesis has aimed to give an account of the various cycles that take place with respect to MTT, marginalization, and madness. It has shown how individuals can be made to feel *stuck in time* in three ways:

(i) madness → impairments to MTT → marginalization
(ii) marginalization → impairments to MTT → madness
(iii) impairments to MTT → madness → marginalization

The first cycle involves a person living with mental illness, suffering with impairments to MTT, and subsequently being socially marginalized as a result. The example I used in the thesis to explore this cycle was my own experience of PTSD. The second cycle beings when one is in an already marginalized group, suffers with impairments to MTT as a result of this, and is at higher risk of being deemed mad as a result. The example I used in this thesis to explore this cycle was the impact racialization has on one's ability to MTT and the psychiatric risks one runs when one disassociates from the collective memory in their social milieu. The final cycle begins in conditions where an individual has suffered impairments to MTT, often by a neurological disorder, which causes them to be deemed mad. This often results in severe marginalization because individuals in the third cycle are particularly vulnerable. The example I used to explore this cycle was that of individuals living with dementia.

I argued that all three cycles result in a feeling of being stuck in time, where one feels mentally unable to move past certain experiences, events, or beliefs. This significantly broadens the scope of what feeling stuck in time is, as the current literature has only discussed individuals with severe memory impairments as qualifying as individuals who feel stuck in time. This section of the thesis will consolidate the personal and relational harms of this experience of feeling stuck in time, which can be drawn out from the articles in the thesis.

The Personal Harms of Feeling Stuck in Time

Disruptions to Self-Trust

As demonstrated in all the chapters of the thesis, feeling stuck in time causes severe disruptions to important self-regarding attitudes such as self-trust. This thesis has shown that self-trust enables us to have positive conceptions of ourselves and our abilities. Without it, individuals lack the ability to successfully MTT and are thereby limited in both their memories of the past and their imaginings in the future. Psychological conditions such as dementia threaten this capacity by disrupting the cognitive mechanisms at play. This can cause individuals to become unsure of themselves and their abilities. Trauma disrupts these cognitive mechanisms in a less obvious way, by dividing one’s attention and effectively splitting one’s sense of self, making it difficult to know which self is competent to rely on. It also affects the affective aspect of self-trust: having a positive conception of oneself. Social conditions, such as racialization and marginalization, disrupt self-trust in a similar way. They do so by ‘splitting’ the self and also by locking individuals into negative collective stereotypes which limit their psychological and social movement in time.

Myopic Imaginal Possibilities

Feeling stuck in time creates a sense of myopic imaginal possibilities. I use the phrase “myopic” in the sense that they are heavily constrained and limited. Philosophers have explored what can occur if someone has an overly open sense of imaginal and epistemic possibility. The result being that this paves the way for delusional thinking (Parrott, 2016). However, philosophers have spent less time exploring the reverse, as to what happens when one has a limited sense of imaginal possibilities. Philosophers exploring the phenomenology of depression have suggested that individuals with depression also suffer with a myopic sense of imaginal possibilities (Broome, 2005; Ratcliffe, 2015). My thesis suggests that more attention needs to be given to this grave harm, as having myopic imaginal possibilities limits one’s very hope for the future, a hope one must have to overcome one’s sense of marginalization (Stockdale, 2021). It also impairs one’s ability to successfully MTT, as demonstrated in Chapter 1 of the thesis. This is because individuals in this situation are not able to conjure up as many simulation episodes as subjects who have an expansive sense of imaginal possibilities.

MTT relies on data from the past and the present to simulate different future episodes (Michaelian, 2016). Feeling stuck in time in either a psychological or social sense limits this data substantially. If one feeling socially stuck in time one’s ability to MTT may be intact but it may be using data from negative collective memories to run, which could result in harmful negative self-perception and internalizing false beliefs about what one is capable of. If one feels psychologically stuck in time, one’s capacity for MTT may begin to malfunction, albeit in less severe ways than currently fleshed out in the literature. Thus, the data MTT requires to be successful is limited by psychological conditions which cause individuals to become highly attentive to negative outcomes and negative perceptions of themselves. It is also limited by social stereotypes that individuals feel they cannot transcend.

The Unsafe Body

Feeling stuck in time affects one's body by making it feel unsafe. Psychological conditions such as trauma do this by keeping an individual in fight-or-flight mode. Social conditions such as racialization do this by conditioning the body's micro-movements and creating a body of violence, as noted in Chapter 2. This occurs because the body wants to resist the marginalization, but the social sphere takes such defiance as a signal for moral failing rather than a sign of moral virtue (Potter, 2016). Neurological conditions such as dementia achieve this by causing an individual to be confused about what they know, whom they love, and what they can do. Marginalized bodies are thus unsafe because if one allows one's body to be defiant, one runs the risk of paying the ultimate social cost.

This harm is particularly insidious because the body and the mind are closely interconnected, such that the body cannot feel unsafe without this feeling affecting the mind. As shown in Chapter 1, feeling unsafe consumes one's attention in a personally harmful way because, without one's faculty of attention, MTT becomes challenging to engage in.

The Relational Harms of Feeling Stuck in Time

Resistance Starts a Cycle of Further Marginalization and Madness

One can attempt to resist the personal harms of being stuck in time, particularly the sense of myopic imaginal possibilities, by defying social stereotypes. Such resistance, however, comes at a price. The price of such resistance has historically been quite high (Maines, 1999; Widdows, 2018; Garcia, 2021). Deviance of any kind results in social suspicion and, in extreme cases social ostracization. Take, for instance, the historical ways in which individuals who differ from the sexual norm resist and the price for such resistance (Foucault, 1988; Rashed, 2019). One can open up one's imaginal possibilities, but MTT remains constrained insofar as such resistance is socially frowned upon. This fact calls our attention to how imaginal possibilities are heavily influenced and dependent on what is socially acceptable. Our relations can work with us to expand our sense of imaginal possibilities or work to act against it.

Blunted Affectivity Can Affect Relationships

One of the most severe relational harms that arise from an individual feeling stuck in time is that the blunted affectivity they suffer has consequences for their relationships. This is demonstrated most starkly in the example of individuals living with dementia, who may forget the names and faces of those they love. Although much less severe, it is present in all cases of feeling stuck in time, as this phenomenology can result in blunted affectivity, such that one either numbs their emotions or becomes resistant to showing them for fear of further marginalization. This, in turn, can compromise one's close connections and put a strain on relationship building. Relating often requires empathizing, and this, more often than not, requires one to successfully MTT. Individuals who feel stuck in time may retain their ability to do so, or they may be impaired. The result is that the individual’s relationships may suffer by being unable to move in time.

Original Contribution to Knowledge

The most original contribution of the thesis is to show that MTT is not an individualistic capacity as previously theorized. It is, instead, heavily socially and relationally mediated. Our social circles can affect both our ability to remember and imagine different kinds of futures by causing individuals to *feel stuck in time*. The notion of feeling stuck in time has had a very narrow focus given the individualistic lens through which MTT is viewed. As MTT is considered an all-or-nothing capacity that one either lacks or has capacity for, individuals only qualify as feeling stuck in time if they lack the ability for MTT altogether. This thesis suggests that this is an error and argues that we need to broaden the range of cases of philosophical interest concerning MTT. I present trauma as a paradigm case where an individual's ability to MTT is impaired but still present. Trauma is theoretically stimulating because it shows how the outside world affects the brain's faculties and in the precise ways it does so. The upshot of this argument that MTT is relationally and socially mediated is significant because it grounds the idea that there are social and relational threats to one's ability to MTT. This is a novel suggestion in both areas of psychiatry and the philosophy of memory.

Continuist theories of MTT rely quite heavily on one having a properly functioning episodic construction system, which reliably and accurately mixes remembering with imagining. This is a central feature of Michaelian's (2016; 2021) account in particular, yet he does not give an account of what a properly functioning episodic construction system is, nor an account of what the function of the episodic construction system is. By examining how the system that produces MTT can malfunction, this thesis aims to give a clearer picture of both what a properly functioning episodic construction system ought to be able to do and what impaired systems are unable to do.

Attention is given to the temporal aspect of specific instances of mental illness, such as depression (Broome, 2005; Ratcliffe, 2015). Yet how interferences in MTT can result in mental distress and marginalization continues to be absent from these discussions. One reason for this is that feminist philosophers, philosophers of psychiatry, and philosophers of memory are currently not in conversation regarding their work. My thesis shows that a dialogue between all three is needed, given the connections between these bodies of work. Previous work by Campbell (2014) began to connect feminist philosophy and the philosophy of memory, as well as research by Black feminist philosophers on processes of forgetting (Allen, 1997), and research by feminist philosophers on imagination (Mackenzie, 2000; Lloyd, 2000). This has subsequently been expanded upon by authors combining feminist insights to Holocaust studies (Hirsch, 2012), as well as feminist work on imagination as a skill (Kind, 2020). The work in this thesis is an extension of this earlier work, by applying these insights to the debate on MTT. It extends this earlier work by showing the cycles individuals with impairments, of any degree, of their ability MTT can suffer from. It suggests that because of the ubiquitous social and relational threats to MTT, psychiatrists need to pay more attention to how memory and mental illness touch.

Finally, impairments to memory are not conceived as clinically compelling at present, with the exception of the large body of psychiatric literature on trauma which does engage with several themes of the thesis itself (Kirmayer, 1996; Aupperle, 2012). Issues with one's memory, outside of this exception, are left to one side in the diagnosis and in the treatment of other mental disorders, as demonstrated in Chapter 2 of this thesis. This thesis suggests that this is a psychiatric oversight and a problem for our understanding of how mental distress occurs. This thesis demonstrates that memory, and memory impairments, in particular, ought to be at the center of our discussions on mental distress, not an afterthought. Impairments to memory matter: to narrative, to one’s self-regarding attitudes, to one’s general hope for the future. This thesis adds to the demand that they are assigned clinical importance too.

Final Conclusion and Summary

This thesis aimed to explore the connections between madness, marginalization, and MTT, specifically by examining what happens when an individual feels psychologically or socially "stuck in time". I sought to show that feeling stuck in time does not simply result from severe memory impairments, as is standardly assumed in the philosophical literature. This thesis demonstrated that this experience is far more ubiquitous than previously assumed and can arise through a mix of both psychological and social conditions. This thesis, therefore, contributes two novel suggestions to the literature in the philosophy of memory. First, that being stuck in time does not only result from neurological disorders but other mental illnesses. Second, that MTT is a relationally and socially mediated capacity. The upshot is that this faculty can be threatened by social and relational factors.

I outlined that the ubiquitous sense of feeling stuck in time results in both personal and relational harms. The personal harms include, but are not limited to: myopic imaginal possibilities, disruptions to self-trust, and feeling unsafe in one's body. The relational harms include, but are not limited to: blunted affectivity in relationships, resisting the myopic possibilities can result in further marginalization.

The arguments in this thesis are significant because discussions which touch on MTT, marginalization, and madness are currently absent from both the literature on the philosophy of memory and the literature on the philosophy of psychiatry. This thesis has argued that memory cannot be placed to one side when discussing mental distress: memory is at the centre of mental health and mental illness.

Implications for Future Research

This thesis opens up various exciting areas for future research. The first area it opens up concerns what the relationship is between MTT impairments and mental distress. I explored what happens when individuals with psychological conditions such as trauma and dementia face impairments to MTT, but research is needed to show whether MTT impairments are present in other mental illnesses. More research will also be needed to show exactly how different psychotherapies can help to repair these impairments for the individual who faces them.

Additional research is also needed to explore the relationship between madness, self-regarding attitudes, and MTT. I have established that self-trust in particular is essential to successful MTT. However, this opens up the question as to whether other self-regarding attitudes, such as self-respect, and self-esteem, are also essential to MTT. It would be fruitful to conduct qualitative research on the intersection between madness, self-regarding attitudes, and one's ability to MTT.

Finally, further research is required to explore how to overcome the dual condition where one psychologically and socially feels stuck in time as this condition is accompanied with the gravest harms. This kind of feeling stuck in time combines both the personal and relational harms. It therefore demands more detailed attention by feminist philosophers, philosophers of memory, and philosophers of psychiatry alike.

Bibliography

Al-Saji, A. (2013). Too Late: Racialized Time and the Closure of the Past. *Insights* 6 (5): 1–13.

Al-Saji, A. (2021). Too Late: Fanon, the Dismembered Past, and a Phenomenology of Racialized Time. In (Eds.) L. Laubscher, D. Hook, & M. U. Desai, *Fanon, Phenomenology and Psychology*, 177–93. New York: Routledge.

Albrecht. G. L., & Devlieger, P. J. (1999). The Disability Paradox: High Quality of Life Against All Odds, *Social Science and Medicine,* 48, 977-988.

Allen, A. (1997). Forgetting Yourself. In (Eds.) D. T. Meyers, *Feminists Rethink the Self.* Oxford: Taylor & Francis.

American Psychiatric Association, (2013). *Diagnostic and Statistical Manual of Mental Disorders [5th Edition],* Arlington, VA: American Psychiatric Publishing.

Applebaum, P. S. (2007). Assessment of Patients' Competence to Consent to Treatment, *New England Journal of Medicine*, 357 (18), 1834-1840.

Atkins, K., & C. Mackenzie. (2008). *Practical Identity and Narrative Agency*. UK: Routledge.

Aupperle, R. L., Melrose, A. J., Stein, M. B., & Paulus, M. P. (2012). Executive Function and PTSD: Disengaging from Trauma, *Neuropharmacology*, 62 (2), 686–694.

Baier, A. (1985). *Postures of the Mind: Essays on Mind and Morals.* Minneapolis: University of Minnesota Press.

Baier, A. (1986). Trust and Antitrust. *Ethics,* 96 (2): 231–60.

Behnke, E. A. (1997). Ghost Gestures: Phenomenological Investigations of Bodily Micromovements and Their Intercorporeal Implications. *Human Studies,* 20 (2), 181-201.

Bloch, D. (2007). *Aristotle on Memory and Recollection: Text, Translation, Interpretation, and Reception in Western Scholasticism*. Leiden: Brill.

Blum, L. (2022). *Iris Murdoch.* Stanford Encyclopedia of Philosophy. Retrieved 17th 2022, from https://plato.stanford.edu/entries/murdoch/

Bok, S. (1978). *Lying: Moral Choice in Public and Private Life*. New York: Pantheon Books.

Boyer, P., & Wertsch, J. (2009). *Memory in Mind and Culture.* Cambridge: Cambridge University Press.

Breslau, N., Davis, G. C., & Peterson, E. (1991). Traumatic Events and Posttraumatic Stress Disorder in an Urban Population of Young Adults. *Archives of General Psychiatry*, 48 (3), 216–222.

Brison, S. J. (2002). *Aftermath: Violence and the Remaking of a Self.* USA: Princeton University Press.

Brison, S. J. (2017). Personal Identity and Relational Selves. In (Eds.) A. Garry, S. J. Khader, A. Stone, *The Routledge Companion to Feminist Philosophy*. UK: Routledge.

Brison, S. J. (2021). Valuing the Lives of People with Profound Intellectual Disabilities. *Philosophical Topics*, 49 (1), 99–122.

Broome, M. (2005). Suffering and the Eternal Recurrence of the Same: The Neuroscience, Psychopathology, and Philosophy of Time. *Philosophy, Psychiatry, & Psychology*, 12 (3), 187–94.

Brothers, D. (1982). Trust Disturbances in Rape and Incest Victims. [Doctoral Thesis: Yeshiva University]. New York: Yeshiva University.

Brown, S. D., & Reavey, P. (2015). *Vital Memory and Affect: Living with a Difficult Past.* New York: Routledge.

Campbell, R., Dworkin, E. R., & Cabral, G. (2009). An Ecological Model of the Impact of Sexual Assault on Women's Mental Health. *Trauma, Violence, & Abuse*, 10 (3), 225–246.

Campbell, S. (2014). *Our Faithfulness to the Past: The Ethics and Politics of Memory, Ed. C.M. Koggel & R. Jacobsen,* Oxford: Oxford University Press.

Carlat, D. J. (2017). *The Psychiatric Interview.* Philadelphia, USA: Wolters Kluwer.

Clark, A., L. Phillipson, & R. Ward. (2021). *Dementia and Place*. Bristol: Bristol University Press.

Colombetti, G., & J. Krueger. (2015). Scaffoldings of the Affective Mind. *Philosophical Psychology*, 28 (8): 1157–76.

Crisp, R. (2018). *The Oxford Handbook of Reasons and Normativity, Ed. D. Star*, Oxford, UK: Oxford University Press.

Damasio, A. R., Tranel, D., & Damasio, H. (1991). Somatic Markers and the Guidance of Behaviour: Theory and Preliminary Testing. In (Eds.) A. L. Benton, H. S. Levin, & H.M. Elsenberg, *Frontal Lobe Function and Dysfunction*, 342–350. Oxford: Oxford University Press.

de Boer, M. E., Hertogh, C., Dröes, R., & Jonker, C. (2010). Advance Directives in Dementia: Issues of Validity and Effectiveness. *International Psychogeriatrics*, 22 (1), 201-208.

Debus, D. (2014). ‘Mental Time Travel’: Remembering the Past, Imagining the Future, and the Particularity of Events. *Review of Philosophy & Psychology,* 5 (3): 333–50.

DeGrazia, D. (1999). Advance Directives, Dementia, and 'The Someone Else Problem'. *Bioethics*, 13 (5), 373-391.

Dillon, R. (1997). Self-Respect: Moral, Emotional, Political. *Ethics*, 107 (2), 226–249.

Draaisma, D. (2000). *Metaphors of Memory: A History of Ideas About the Mind.* Cambridge: Cambridge University Press.

Drabinski, J. E. (2013). Fanon's Two Memories. *South Atlantic Quarterly*, 112 (1), 5-22.

Du Bois, W. E. B. (2018). *The Souls of Black Folk, Ed. P.H. Hinchey*. Gorham ME: Myers Education Press.

Dworkin, R. M. (1994). *Life's Dominion: Argument About Abortion and Euthanasia*. London: Harper Collins Publishers.

Eyre, R. (Director). (2001). *Iris [film]*. BBC Films: Fox Iris Productions.

Fanon, F. (1961). *Wretched of the Earth,* R Philcox (trans.), Grove Press: New York.

Fanon, F. (1964). *Alienation and Freedom, Ed. J. Khalfa and R. Young,* S. Corcoran (trans.), New York: Bloomsbury Press.

Fanon, F. (1967a). *Black Skin, White Masks,* C Markmann (trans.), New York: Grove Press.

Fanon, F. (1967b). *Toward the African Revolution, Political Essays,* H. Chevalier (trans.), New York: Monthly Review Press.

Fanon, F. (1975). *Pour La Révolution Africaine : Écrits Politiques*. Paris : F. Maspero.

Fazel, S., Hope, T. & Jacoby, R. (1999). Dementia, Intelligence, and the Competence to Complete Advance Directives. *The Lancet*, 354, 9172.

Fehige, C. and Wessels, U. (2011). *Preferences.* Berlin: De Gruyter.

Fellows, L. (2015). Competency and Consent in Dementia, *Journal of the American Geriatrics Society*, 46 (7), 922-926.

Fernando, S. (2017). *Institutional Racism in Psychiatry and Clinical Psychology: Race Matters in Mental Health,* Cambridge: Palgrave MacMillan.

Fins, J. J. (2017). Mosaic Decision-making and Reemergent Agency After Severe Brain Injury. *Cambridge Quarterly of Healthcare Ethics*, 27 (1), 163–74.

Fivush, R., & Graci, M. (2017). Memory and Social Identity. In (Eds.) S. Bernecker & K. Michaelian, *The Routledge Handbook of Philosophy of Memory*. UK: Routledge.

Fluchtmann, J., Glenny, A. M., Harmon, N., & Maibom, J. (2021). The Gender Application Gap: Do Men and Women Apply for the Same Jobs? *IZA Discussion Papers*, 14906.

Foucault, M. (1988). *Madness and Civilization: A History of Insanity in the Age of Reason*. UK: Vintage Books.

Fraassen, B. C. V. (1995). Belief and the Problem of Ulysses and the Sirens, *Philosophical Studies*, 77, 7-37.

Freyd, J. J. (1996). *Betrayal Trauma: The Logic of Forgetting Childhood Abuse*. Cambridge: Harvard University Press.

Fried, T., Knight, S., McMahan, R., & Sudore, R. (2013). Advance Care Planning Beyond Advance Directives: Perspectives from Patients and Surrogates, *Journal of Pain and Symptom Management*, 46 (3), 355-365.

Gaetz, S., E. Dej, T Richter, & M. Redman. (2016). *The State of Homelessness in Canada*. Toronto, Canada: Canadian Observatory Press.

Gale, S. A., D. Acar, and K. R. Daffner. (2018). Dementia. *American Journal of Medicine*, 1161–69.

Garcia, M. (2018). *On Ne Naît Pas Soumise, On Le Devient*. France : Flammarion.

Gerrans, P, and J. Kennett. (2017). Mental Time Travel, Dynamic Evaluation, and Moral Agency. *Mind*, 126 (501), 259–68.

Gerrans, P. (2018). Painful Memories. In (Eds.) D. Debus, D. Perrin, & K. Michaelian, *New Directions in Philosophy of Memory*. UK: Taylor & Francis.

Gerrans, P., & Kennett, J. (2010). Neurosentimentalism and Moral Agency. *Mind*, 119, 475.

Gerrans, P., & Kennett, J. (2017). Mental Time Travel, Dynamic Evaluation, and Moral Agency. *Mind*, 126 (501), 259–268.

Gerrans, P., & Sander, D. (2014). Feeling the Future: Prospects for a Theory of Implicit Prospection. *Biology & Philosophy*, 29 (5), 699–710.

Gilmoor, A., Vallath, S., Regeer, B., & Bunders, J. (2020). “If somebody could just understand what I am going through, it would make all the difference”: Conceptualizatons of Trauma in Homeless Populations Experiencing Severe Mental Illness. *Transcultural Psychiatry*, 57 (3), 455–467.

Giordano, C. (2020). *Boris Johnson Says Attacking Statues is “Lying About Our History” and Protests Have Been “Hijacked by Extremists”,* The Independent, Retrieved 10th June, 2022 from: https://www.independent.co.uk/news/uk/politics/boris-johnson-statues-churchill-mandela-colston-protests-black-lives-matter-a9562626.html

Goering, S. (2020). Why Should Adamancy of an Uninformed View Give Moral Weight? *The American Journal of Bioethics*, 20 (8), 78–79.

Goffman, E. (1991). *Asylums: Essays on the Social Situation of Mental Patients and Other Inmates,* London: Penguin.

Gold, J., & Gold., I. (2015). *Suspicious Minds: How Culture Shapes Madness*. USA: Simon and Schuster.

Govier, T. (1993). Self-Trust, Autonomy, and Self Esteem. *Hypatia*, 8 (1), 99–120.

Groleau, D., A. Young, & L. J. Kirmayer. (2006). The McGill Illness Narrative Interview (MINI): An Interview Schedule to Elicit Meanings and Modes of Reasoning Related to Illness Experience. *Transcultural Psychiatry,* 43 (4), 671–91.

Groll, D. (2015) Medicine and Well-being. In (Eds.) G. Fletcher, *The Routledge Handbook of Philosophy of Well-Being*, London: Routledge.

Hacking, I. (1995). *Rewriting the Soul: Multiple Personality and the Sciences of Memory*. Princeton University Press: New Jersey.

Hacking, I. (1999). *The Social Construction of What?* USA: Harvard University Press.

Halbwachs, M. (1980). *The Collective Memory.* New York, NY: Harper & Row.

Halbwachs, M. (1992). *On Collective Memory.* L. Coser (trans.), Chicago, IL: University of Chicago Press.

Harewood, D. (2021) *Maybe I Don't Belong Here: A Memoir of Race, Identity, Breakdown and Recovery,* Bluebird: UK.

Harewood, D. (2021). 'I came close to death': David Harewood on Racism and Psychosis, The Guardian. Retrieved on 10th June 2022, from https://www.theguardian.com/culture/2021/aug/21/david-harewood-homeland-racism-psychosis-close-to-death

Hawkins, J. (2014). Well-Being, Time, and Dementia, *Ethics,* 124 (3), 507-542.

Hirsch, M. (2012). The Generation of Post-Memory: Writing and Visual Culture After the Holocaust. New York: Columbia University Press.

Hope, T. (1994). Personal Identity and Psychiatric Illness. *Royal Institute of Philosophy Supplement*, 37, 131-143.

Hope, T., Slowther, A. and Eccles, J. (2005). Best Interests, Dementia and the Mental Capacity Act. *Journal of Medical Ethics*, 35 (12).

Irish, M. (2016). Semantic Memory as Essential for Future-Orientated Mental Time Travel. In (Eds.) K. Michaelian, S. B. Klein, K. K. Szpunar, *Seeing the Future.* Oxford: Oxford University Press.

Isurin, L. (2017). *Collective Remembering: Memory in the World and in the Mind.* Cambridge, UK: Cambridge University Press.

Jaworska, A. (1999) Respecting the Margins of Agency: Alzheimer’s Patients and The Capacity to Value. *Public Affairs*, 28(2), 105–138.

Jecker, N.S. (2020). *Ending Midlife Bias*. New York: Oxford University Press.

Jones, K. (1996). Trust as an Affective Attitude. *Ethics* 107 (1): 4–25.

Jones. K. (2012). Trustworthiness. *Ethics* 123 (1): 61–85.

Kennett, J. (2009). Mental Disorder, Moral Agency, and the Self. In (Eds.) B. Steinbock, *The Oxford Handbook of Bioethics*. Oxford: Oxford University Press.

Kind, A. (2020). The Skill of Imagination. In (Eds.) E. Fridland & C. Pavese, *The Routledge Handbook of Skill and Expertise*, Oxford: Routledge.

Kirmayer, L. (1996) *Landscapes of Memory: Trauma, Narrative, and Dissociation*. In (Eds.) P. Antze & M. Lambek, *Tense Past: Cultural Essays in Trauma and Memory*, 173-199, Routledge: London.

Kirmayer, L. (2012). Rethinking Cultural Competence. *Transcultural Psychiatry*, 49 (2), 149-164.

Kirmayer, L., Felipe, A., & Walsh, E. (2022). The Social Determinants of Health in Neuroscience: An Overview. *Transcultural Psychiatry*, *forthcoming*.

Kittay, E. (2010). "The Personal is Political is Philosophical: A Philosopher and Mother of a Cognitively Disabled Person Sends Notes from the Battlefield." In (Eds.) E. F. Kittay & L. Carlson, *Cognitive Disability and Its Challenge to Moral Philosophy,* 393-413. Malden, MA: Wiley-Blackwell.

Kitzinger, J. and Kitzinger, C. (2016). PPIW Report Increasing the Awareness and Uptake of Advance Directives. Retrieved June 17th, 2022, from http://ppiw.org.uk/files/2016/02/PPIW-Report-Increasing-the-awareness-and-uptake-of-Advance-Decisions.pdf

Korsgaard, C. (1988). Personal Identity and the Unity of Agency: A Kantian Response to Parfit. *Philosophy & Public Affairs*, 18, 101–32.

Kozlowska, K. (2007). Intergenerational Processes, Attachment, and Unexplained Medical Symptoms. *Australian and New Zealand Journal of Family Therapy*, 28 (2), 88-99.

Lambek, M. (1996). The Past Imperfect: Remembering as Moral Practice. In (Eds.) P. Antze & M. Lambek, *Tense Past: Cultural Essays in Trauma and Memory*, 235-255, Routledge: London.

Langenbacher, E., Shain., Y., & Ebrary, I. (2010). *Power and the Past: Collective Memory and International Relations.* USA: Georgetown University Press.

León, F., Szanto, T., & Zahavi, D. (2019). Emotional Sharing and the Extended Mind, *Synthese*, 196 (12), 4847 – 4867.

Levy, B. R., Zonderman, A. B., Slade., M. D., and Ferrucci, L. (2009). Age Stereotypes Held Earlier in Life Predict Cardiovascular Events Later in Life. *Psychological Science*, 20 (3), 296–98.

Lindemann, H. (2009). Holding One Another (Well, Wrongly, Clumsily) in a Time of Dementia. *Metaphilosophy,* 40 (3-4), 416-424.

Lindemann, H. (2014). What and When to Let Go: Identities at the End of Life. In (Eds.) H. Lindemann*, Holding and Letting Go: The Social Practice of Personal Identity*. Oxford: Oxford University Press.

Lloyd, G. (2000). Individuals, Responsibility, and the Philosophical Imagination. In (Eds.) N. Stoljar & C. Mackenzie, *Relational Autonomy,* 112-123. Oxford: Oxford University Press.

Mackenzie, C., & Stoljar. N. (2000). *Relational Autonomy: Feminist Perspectives on Autonomy, Agency and the Social Self.* Oxford: Oxford University Press.

Mackenzie, C. (2000). Imagining Oneself Otherwise. In (Eds.) N. Stoljar & C. Mackenzie, *Relational Autonomy,* 124-150. Oxford: Oxford University Press.

Mackenzie, C. (2009). Personal Identity, Narrative Integration and Embodiment. In (Eds.) S. Campbell, L. Meynell, & S. Sherwin, *Embodiment and Agency*, 100–125. University Park, PA: The Pennsylvania State University Press.

Mackenzie, C., & Poltera, J. (2010). Narrative Integration, Fragmented Selves, and Autonomy. *Hypatia*, 25 (1), 31-54.

Maines, R. P. (1999). *"Hysteria", The Vibrator, and Women's Sexual Satisfaction: The Technology of Orgasm*. Baltimore, USA: Johns Hopkins Press.

Martin, P., and A. Rowe. (2010). *Iris Murdoch: A Literary Life*. London: Palgrave Macmillan.

McGeer, V., and P. Pettit. (2017). The Empowering Theory of Trust. In (Eds.) P. Faulkner & T. Simpson, *The Philosophy of Trust*. Oxford: Oxford University Press.

McLeod, C, and S. Sherwin. (2000). Relational Autonomy, Self-Trust, and Health Care for Patients Who Are Oppressed. In (Eds.) N. Stoljar and C. Mackenzie, *Relational Autonomy*, 259–279, Oxford: Oxford University Press.

McMahan, J. (2003). *The Ethics of Killing: Problems at the Margins of Life.* New York: Oxford University Press.

Meyers, J. (2013). *Remembering Iris Murdoch: Letters and Interviews*. New York: Palgrave Macmillan.

Michaelian, K. (2016). *Mental Time Travel*. Cambridge, MA: MIT Press.

Michaelian, K. (2020). Confabulating as Unreliable Imagining: In Defence of the Simulationist Account of Unsuccessful Remembering. *Topoi*, 39 (1), 33–148.

Michaelian, K. (2021). Imagining the Past Reliably and Unreliably: Towards a Virtue Theory of Memory. *Synthese*, 199 (3–4), 7477–7507.

Michaelian, K., & Sutton, J. (2017). *Memory*. The Stanford Encyclopaedia of *Philosophy. Retrieved on 10th June, 2022 from https://plato.stanford.edu/archives/sum2017/entries/memory/>*

Michaelian, K., & Sutton, J. (2019). Collective Mental Time Travel: Remembering the Past and Imagining the Future Together. *Synthese*, 196, 4933–4960.

Miller, C. (2020). *Know My Name; A Memoir*. London: Penguin.

Mingon, M., & Sutton, J. (2021) Why Robots Can't Haka: Skilled Performance and Embodied Knowledge in the Mãori Haka, *Synthese*, 199, 4337-4365.

Mitchell, W. (2018). *Somebody I Used to Know.* London: Bloomsbury Publishing

Motta, V. (2021). Key Concept: Loneliness. *Philosophy, Psychiatry, & Psychology*, 28 (1), 71–81.

Moye, J. and Marson, D. C. (2007). Assessment of Decision-Making Capacity in Older Adults: An Emerging Area of Practice and Research. *The Journals of Gerontology*, 62 (1), 3-11.

Mukandi, B. (2019). *The North African Syndrome: Traversing the Distance to the Cultural "Other*". In (Ed.) S. Tekin & R. Bluhm, *Bloomsbury Companion to the Philosophy of Psychiatry*, 413-428, New York, NY: Bloomsbury Academic.

Murdoch, Iris. (2003). *Metaphysics as a Guide to Morals*. Vintage Classics. London: Vintage Books.

Narváez, R. F. (2013). *Embodied Collective Memory: The Making and Unmaking of Human Nature.* USA: University Press America.

Novick, P. (1999). *The Holocaust in American Life.* USA: Mariner books.

Nozick, R. (2001). *Anarchy, State, and Utopia*. Oxford: Blackwell Publishing.

Okpaku, S. (1998). *Clinical Methods in Transcultural Psychiatry*. USA: American Psychiatric Association Publishing.

Olson, E. T. (2019). *Personal identity*. Stanford Encyclopedia of Philosophy. Retrieved June 17, 2022, from https://plato.stanford.edu/entries/identity-personal/

Parfit, D. (1971). Personal Identity. *The Philosophical Review*, 80 (1): 3–27.

Parrot, M. (2016). Bayesian Models, Delusional Beliefs, and Epistemic Possibilities, *The British Journal of Philosophy of Science*, 67 (1), 271-296.

Paul, L. A. (2016) *Transformative Experience*. Oxford: Oxford University Press.

Perrin, D., & Michaelian, K. (2017). Memory as Mental Time Travel. In (Eds.) S. Bernecker & K. Michaelian, *The Routledge Handbook of Philosophy of Memory*. UK: Routledge.

Persad, G. (2019) Authority Without Identity: Defending Advance Directives via Posthumous Rights Over One's Body, *Journal of Medical Ethics;* 45, 249-256.

Pittenger C. (2013). Disorders of Memory and Plasticity in Psychiatric Disease, *Dialogues Clinical Neuroscience*, 15 (4), 455-63.

Potter, N. (2016). *The Virtue of Psychiatric Defiance*. Oxford: Oxford University Press.

Rashed, M. A. (2019). *Madness and the Demand for Recognition: A Philosophical Inquiry into Identity and Mental Health Activism.* Oxford: Oxford University Press.

Ratcliffe, Matthew. (2015). *Experiences of Depression: A Study in Phenomenology*. Oxford, UK: Oxford University Press.

Reiland, R. (2004). *Get Me Out of Here: My Recovery from Borderline Personality Disorder*. Centre City, MN: Hazelden Foundation.

Ricœur, Paul. (2003). *La Mémoire, L'Histoire, L'Oubli*. Points Série Essais 494. Paris: du Seuil.

Rothberg, M. (2009). *Multidirectional Memory: Remembering the Holocaust in the Age of Decolonization.* Stanford: Stanford University Press.

Saks, E. R. (2007). *The Center Cannot Hold: My Journey Through Madness*. New York, USA: Hyperion

Scheman, N. (1983). Individualism and the Objects of Psychology. In (Eds.) S. Harding & M. B. Hintikka, *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology, and Philosophy of Science*, 225–44. Dordrecht: Reidel.

Sehgal, A., Galbraith, A., Chesney, M., Schoenfeld, P., Charles, G., & Lo, B. (1991) How Strictly Do Dialysis Patients Want Their Advance Directives Followed? *Journal of the American Medical Association*, 267 (1), 59-63.

Shiffrin, S. (2004) Autonomy, Beneficence, and the Permanently Demented. In (Eds.) J. Burley, *Ronald Dworkin and His Critics*, 195-217. New Jersey: Blackwell Publishing.

Shinn, M, and J. Khadduri. 2020. *In the Midst of Plenty: Homelessness and What to Do About It*. USA: Wiley Blackwell.

Small, S.A. (2021). *Forgetting: The Benefits of Not Remembering*. New York: Crown Publishing.

Sokolowski, M. (2018). *Dementia and the Advance Directive – Lessons from the Bedside.* USA: Springer.

Sterelny, K. (2010). Minds: Extended or Scaffolded? *Phenomenology and Cognitive Sciences*, 9, 465–481.

Stockdale, K. (2021). *Hope Under Oppression.* New York, USA: Oxford University Press.

Strawson, G. (2004). Against Narrativity. *Ratio*, 17 (4), 428–452.

Suddendorf, T., & Corballis, M. C. (1997). Mental Time Travel and the Evolution of the Human Mind. *Genetic, Social, and General Psychology Monographs*, 123 (2), 133–167.

Suddendorf, T., & Corballis, M. C. (2007). The Evolution of Foresight: What is Mental Time Travel and is it Unique to Humans? *Behavioural and Brain Sciences*, 30 (3), 299-313.

Suddendorf, T., Addis, D. R., & Corballis, M. C. (2009). Mental Time Travel & the Shaping of the Human Mind. *Philosophical Transactions of the Royal Society*, 364, 1317–1324.

Sutton, J. (2018). Shared Remembering and Distributed Affect. In (Eds.) K. Michaelian, D. Debus, and D. Perrin, *New Directions in the Philosophy of Memory.* Oxford: Routledge.

The, A., Pasman, R., Onwuteaka-Philipsen, B., Ribbe, M., and Van der Wal, G. (2002) Withholding the Artificial Administration of Fluids and Food from Elderly Patients with Dementia: Ethnographic Study, *The British Medical Journal*, 325, 1326-1330.

Thomson, J. (1971). A Defense of Abortion. *Philosophy and Public Affairs*, 1 (1), 47-66.

Tulving, E. (1983). *Elements of Episodic Memory*. Oxford, UK: Clarendon Press.

Ubel, P., Schwarz, N., Loewenstein, G., and Smith, D. (2005). Misimagining the Unimaginable: The Disability Paradox and Health Care Decision Making. *Health Psychology*, 24 (4), S57-S62.

Van der Kolk, B. A. (2015). *The Body Keeps the Score: Brain, Mind and Body in the Healing of Trauma*. UK: Penguin Books.

Vezzoni, C. (2005). *The Legal Status and Social Practice of Treatment Directives in the Netherlands.* Groningen: RUG.

Wagoner, B. (2017). *The Constructive Mind: Bartlett's Psychology in Reconstruction.* Cambridge: Cambridge University Press.

Walker, M., & Rosen, W.B. (2005). *How Connections Heal: Stories from Relational-Cultural Therapy.* New York: Guilford Press.

Wells, C. (1977). *Dementia*. Pennsylvania, USA: Davis Company.

Widdows, H. (2018). *Perfect Me Beauty as an Ethical Ideal.* New Jersey, USA: Princeton University Press.