www.auc.nl

UNIVERSITY OF AMSTERDAM

VU UNIVERSITY AMSTERDAM

AUC Undergraduate Journal of Liberal Arts & Sciences

Capstone Issue Vol. 10 2018

The AUC Undergraduate Journal of Liberal Arts and Sciences is a biannual, interdisciplinary publication showcasing outstanding undergraduate academic papers. The Journal aims to demonstrate the strength of undergraduate scholarship at AUC, to reflect the intellectual diversity of its academic programme, to encourage best research and writing practices, to facilitate collaboration between students and faculty across the curriculum, and to provide students with opportunities to gain experience in academic reviewing, editing and publishing. The Editorial of the Journal is constituted of members of the InPrint board, a registered AUCSA committee.

Foreword

In the semester prior to graduation, every AUC student is required to carry out an independent research project within their intended major (Sciences, Social Sciences or Humanities), referred to as the Capstone. This project is meant to have students engage with the current academic dialogue within their fields. With this year'€™s Capstone Issue, which includes six capstones across all three majors from students from the class of 2018, InPrint showcases the ability of AUC students to generate insightful interdisciplinary research.

Additionally, all Capstones published in this issue have undergone a meticulous editing process carried out by the Editorial Board of InPrint to further improve their clarity. A workshop providing the basic tools for this editing process was provided by lecturer and linguist Dr. Lotte Tavecchio. Finally, assistance with the LATEX 2$_\varepsilon$ formatting for this issue was graciously provided by former Editor-in-Chief Phillip Hartout.

We hope that reading this capstone issue gives our readers both insight into the high level of academic achievement of AUC students as well as how the integration of skills and talents achieved through a liberal arts education can come together to create something truly excellent.

Lanie Preston, on behalf of InPrint

Contents

Social Sciences

Alt-Reich: Why We Fail to Recognize Far Right Extremism - Mapping the Alternative Right's Repertoire of Violence

Nasiba Sharipova

Abstract

This study analyses the development of the Alternative Right movement from an agglomeration of online platforms into a violent white nationalist movement, with a focus on this movement in the American context. Drawing on scholarship in media, as well as peace and conflict studies, this thesis conducts visual and discursive analyses of relevant images, events, and texts, specifically looking at the Alternative Right's behavior on the Internet, as well its coverage by traditional news media. The analyses reveal that through narrative building and collective identity construction the movement propagates a worldview in which white people are oppressed, and herein places an implicit call for the collective it has constructed to stand up against the oppressor. The Alternative right's knowledge of the principles of traditional media discourse allows it to manipulate the media in such a way that the violence perpetrated at the hands of its sympathizers is not linked to the movement or its ideology. The overall goal of this thesis is to urge an adequate problematization of Alternative right's progression toward violence.

Keywords and phrases: *Alt-right, political violence, white supremacy, radical right, social media*

Introduction

"Hail Trump, hail our people, hail victory!". Less than two weeks after the victory of the presidential candidate Donald Trump in 2016, Richard B. Spencer chose these words to welcome over 200 attendees to the annual conference of the National Policy Institute. The conference guests applauded their unofficial leader, who in only a matter of months would enjoy international recognition as the face of the Alternative Right, and returned his greeting with chants and Nazi salutes (Goldstein, 2016).

During the 2016 United States presidential election, it was not only Donald Trump who made headlines with controversial statements; the world was confronted with a Frankenstein's monster that had taken the shape of the Alternative Right. With its unconventional practices such as "trolling", America's organically-developed, bottom-up, right-wing political movement left many people puzzled. As a movement, the Alternative Right is a vanguard for radical right-wing ideology and white identity, and with its nativity story tracing back to the dark corners of the Internet (Salazar, 2018). Once only referring to a series of vaguely connected websites where online subcultures and identity politics were prevalent, today the Alt-right not only represents a movement that is highly skilled in conducting politics online, but also a large organization with a wide net of supporters. Relying on legions of white middle class men, the Alternative Right is on a quest to tell America a story: one where America is in a crisis that traditional methods cannot resolve, one where whites are on the edge of extinction, and one where liberalism, feminism and multiculturalism are seen as a threat to the future of America. In the prop-

agation of its narrative, Alt-right unapologetically proposes white and male superiority and opposes both progressive and conservative America (Marwick & Lewis, 2017; Salazar, 2018). Feminism and "political correctness" are rejected, journalists and media are deemed deceitful, and the pervasiveness of inflammatory language across the platforms used by the movement reveals that Alt-right considers the First Amendment as the "right to offend" (Lyons, 2017). Primarily interacting online, the white nationalist movement is often found on platforms such as Reddit, 8chan and 4chan, as well as mainstream social media, where its sympathizers discuss the issues of politics, race and gender. Alt-right is further known for its embracing of memes, which in the online context refers to images paired with a caption representing a common expression of comedic or relatable information (Lyons, 2017; Marwick & Lewis, 2017). The reliance of the Alternative Right on the Internet is strategic, as the medium assists in the process of mass communication, facilitates the construction of a collective identity, allows the movement to disseminate its message faster, requires little to no resources, and often secures anonymity (Adams & Roscigno, 2005; Lyons, 2017). Additionally, the sphere of obscurity, ephemerality, and irony that governs on these platforms allows those previously not interested in politics, to flirt with hateful and often racist ideas, without facing the obligation of accountability. When instrumentalized for political ends, the Internet lends many opportunities to new forms of collective action, such as organized online campaigns or crusades targeting enemies of the movement (Marwick & Lewis, 2017; Neiwert, 2017). Today these are commonplace practices of Alternative Right, and the competency of the movement

to utilize the Internet in the said manner is exactly what makes it unique and frightening.

The worldview painted by the Alternative Right implicitly demands action against its enemies, denouncing democratic methods, and with reference to eugenics, the movement asks its followers to conspire in the creation of a pure, white America (Lyons, 2017). In recent years the United States has been plagued by hate crimes and gun violence, with some perpetrators proudly claiming their violence to be race-fueled; through this we see the radicalized face of the Alternative Right. A report published by the Southern Poverty Law Center, looked at incidents from 2014 to 2018, and estimated the number of deaths linked to the Alternative Right a to be 43 (Amend & Hankes, 2018). Violence was often a result of the successful internalization of the Alt-right's narrative of white marginalization, and a recurring motive in most of the researched cases. Specifically targeting males between the ages of 16 and 24, the Alt-right intentionally prompts feelings of alienation, grievance, and discrimination within the demographic that is most prone to committing mass violence (Amend & Hankes, 2018). Deceived by seemingly harmless imagery like that of the meme Pepe the Frog, and still puzzled about the practice of 'trolling', the public and mainstream media often fail to draw a link between the Alt-right and the violence it orchestrates.

Being one of America's youngest and fastest growing movements, the Alternative Right represents an interesting topic for journalistic, as well as scientific, inquiry. The movement has been under the watch of civil rights advocacy group, the Southern Poverty Law Center, and has been extensively examined by the Data and Society research institute in their 2017 Media Manipulation and Data Disinformation Online report. While the findings help to visualize the physical as well as the virtual presence and legacy of the Alternative Right, an ambition to conceptualize this new face of hate is lacking, leaving the Alt-right's path to violence inadequately problematized. In light of this, the research objective driving this thesis is to conceptualize the dangerous development of the white supremacist movement and investigate how its violent arm has remained under the radar.

This thesis will draw a blueprint of the Alt-right's progression from violent radical discourse to physical violence and radical extremism. In doing so, this study will venture into communication and media scholarship, as well as peace and conflict studies. A guiding assumption of this thesis is that the construction of a collective identity with shared values and grievances within social movements facilitates, and even compels (collective) action. Themes needed to advance this argument will be addressed in three chapters that will make up the body of this project. A literature review referring to leading research in the field will serve as the groundwork for this project and will assist the analysis of the chosen case studies. The first chapter will follow the development of the Alternative

Right from the onset of the Gamergate scandal to the ascending of its front men into the chambers of the White House. This account will be enriched with the examination of Alt-right's online culture and traditions, as well as an overview of its recently established digital sphere of influence. The second chapter will analyze how the Alt-right's investment into the online sphere has transformed the face of white nationalism, specifically looking at practices such as implicit storytelling and narrative building. This will be achieved through conducting a comparative case study analysis of comics disseminated by the White Aryan Resistance between 1980's and 1990's, and articles published on hyperpartisan news website AltRight.com to assess how newer forms of hate speech differ from older ones. Here, this thesis will pay attention to how the Alternative Right constructs cultural frames of interpretation as a semi-sophisticated way of creating propaganda. Finally, Johan Galtung's notions of structural and physical violence will be applied to verify that violent discourse directed by interpretational frames incites and justifies physical violence. The third and final chapter of this thesis will assess the competency of the Alternative Right in manipulating mainstream media by drawing on the expertise acquired during the Gamergate controversy, its appropriation of the practice of "trolling", and its understanding and skillful instrumentalization of concepts like "Poe's Law" and the "Overton Window". The Pizzagate fake news scandal will act as the case study for the final chapter.

Chapter One: The Alternative Right: The Birth and Rise

The lineage of the "Alternative Right" traces back to the first decade of the 21st century. First

coined by Richard Spencer in 2008, the term immediately became synonymous with the vanguard of white interests and identity, appealing to those who no longer felt represented by classical conservatism. This had been previously embodied by the Republican party which was now "too soft", having had had surrendered to the powers of globalization and multiculturalism (Arciniega, 2017; Salazar, 2018). The development of Alt-right into the movement that it is known as today was largely facilitated by the affordances of modern media channels, most importantly the Internet. The Alternative Right's virtual sphere of influence stretches over a multitude of websites like 4chan, 8chan, Reddit, personal blogs, interest-based news bulletin boards, news platforms, and online magazines like Altright.com and the Breitbart News Network (Marwick & Lewis, 2017). Established in 2010 by Spencer, Altright.com served as one of the main platforms for the development of the movement as it embodied a virtual podium for rightist ideology, representing a breeding place for the sympathizers of racism, misogyny, libertarianism, nativism, Islamophobia and other outlooks situated on the far-right spectrum (Lyons, 2017). Initially, the Alternative Right was a classification reserved for an exclusive intellectual circle mostly active on the margins of the Internet, who self-identified with the movement and its values. Soon, the Alt-right expanded its influence beyond its familiar domains like 4chan and Reddit, and started exerting its influence on mainstream online networks (Marwick & Lewis, 2017). This expansion was largely facilitated by the surge of an online dispute that is today known as the "Gamergate" scandal. The online controversy that took the internet by storm in 2013, was an outburst of accumulated tensions within the video-gaming community, and was chiefly concerned with the internal politics and inequalities within this circle (Marwick & Lewis, 2017; Neiwert, 2017). The dispute was heated by the vigorously antagonizing rhetoric adopted by the conservative camp in response to the calls made by the liberal faction for a gender-inclusive gaming environment and female representation in video games. The conservative opposition perceived such demands as a declaration of a "culture war", and called for the resistance against the "politically correct" multiculturalist agenda that the feminist camp was supposedly propagating (Massanari, 2017).

The angry rhetoric spread over to a multitude of online forums, chat rooms, and mainstream community-based platforms, inciting hate towards the liberal faction. Despite the fragmentation of the dispute across various levels of the Internet, a number of individuals found themselves at the center of the scandal. The controversy was initially sparked by the positive reception of a female-friendly game, Depression Quest, developed by feminist game designer Zoe Quinn (Massanari, 2017). Quinn's game situated the player in the shoes of a person suffering from clinical depression and guided them through their hardships. The creation was lauded for its innovative approach toward gaming and for its ability to utilize gaming as a medium that facilitated mutual understanding (Massanari, 2017). It was the affirmative response of this new game that provoked backlash from livid anti-feminist gamers. In the view of these gamers, Depression Quest's feminist and inclusive narrative deviated too much from the traditional, "male" combat-oriented games, and was therefore considered unworthy of its success (Neiwert, 2017). The resentment toward this game quickly spilled over from the virtual into the real world. Zoe Quinn was soon the target of death and rape threats that were sent to her social media accounts, as well as her home, marking the beginning of the Gamergate scandal (Marwick & Lewis, 2017). A year later, in 2014, the online dispute escalated further. Quinn's ex-boyfriend published a story in which he implied that the success of the game was assisted by Quinn's new boyfriend and a video game critic, Nathan Grayson (Marwick & Lewis, 2017). Rapidly, the story had developed a life of its own: the critics of the feminist game maker claimed that her boyfriend had published a positive review of Depression Quest, thus playing a pivotal role in the success of the game (Neiwert, 2017). This review, did in fact not exist, but this did not shift the current. The anti-feminist gamers crowned this supposed breach of journalistic ethics as a major piece of evidence for the rising feminist, liberal and therefore anti-white, anti-cisgendered, and anti-male sentiment within the gaming industry. From this point, the scandal mostly took place under the hashtag #Gamergate, as coined in a Tweet by conservative Hollywood actor, Adam Baldwin (Massanari, 2017; Neiwert, 2017).

Zoe Quinn was not the only woman affected by this scandal. Along with Quinn, video game developer Brianna Wu and video game critic Anita Sar-

keesian fell victim to ruthless harassment. What started off as a defense of so-called "gamer culture" increasingly started to resemble a bullying campaign that eventually forced all three women to flee their homes (Perreault & Vos, 2018). Due to the severity of the developments within this online debacle, and its shift from the virtual into the physical world, 4chan, the website that had initially served as the main host for the debate, announced in September of 2014 to put a ban on any threads discussing "#Gamergate" (Neiwert, 2017). This decision, however, did not reach its desired objective. Instead of putting a halt to the scandal, the announcement prompted the Gamergaters to move to the rival, less curated website 8chan, as well as more mainstream platforms like Twitter and Reddit (Marwick & Lewis, 2017). United across these platforms the Gamergaters developed their exclusive jargon, referring to their victims as "special snowflakes" or "social justice warriors", and mockingly calling their websites "safe spaces" (Neiwert, 2017). This faction of gamers also displayed sympathy towards conspiracy thinking, most practically displayed in the fact that they invoked the immediate threat of "cultural Marxism" to the future of Western civilization as a justification for their online crusade against liberal values. This conspiracy gained wide appeal among the #Gamergate community, and relevant names within the movement advanced the idea that feminists and liberals within the gaming industry were motivated by a longing to sabotage the long-standing civilization of conventional masculine tradition (Marwick & Lewis, 2017). Together with the popularity of this sentiment, a common uniting theme began to root itself among the discussions: the #Gamergaters shared the belief that white men had fallen victim to systematic intolerance and oppression at the hand of left-wing liberals, and dubbed the moderate response from their fellow conservatives as a sign of weakness and subjection to the forces of multiculturalism (Wilson, 2015). As the media gradually lost interest in the scandal, the topics discussed deviated away from gaming toward politics and conspiracies. Now continuing as a movement, the gamers, the online trolls, white nationalists and the sympathizers of conservative values, had expanded their target group from liberals, women and feminists to ethnic minorities, the LGBTQ+ community, Muslims, Jews and moderate conservatives, and coined themselves the "Alternative Right" (Massa-

nari, 2017; Neiwert, 2017).

The events of the #Gamergate scandal have majorly influenced the political and ideological outlook of Alt-right. The group increasingly started to endorse strong anti-female sentiments and investing in establishing the view that women were in need of men to dominate them and incapable of political participation (Marwick & Lewis, 2017). This extreme hostility reflected the views formed during #Gamergate, namely that white men were facing oppression at the hand of women. Further, the rhetoric adopted by the group is often infused with anti-Semitic, anti-LGBTQ+, and anti-immigrant dispositions that point towards Alt-right's endorsement of eugenics as a tool for White supremacy (Lyons, 2017). Put simply, the core values of Alternative Right are rooted in the belief that liberal elites, and the so-called social justice warriors pose a danger to "White identity", Western values and men's rights through their use of "political correctness" (Futrell & Simi, 2017). Important to mention is that along with its hatred for ethnic minorities, liberals and women, the Alternative Right is also extremely hostile towards conservatives who appear tolerant toward some of liberal values, whom the Alt-right refers to as "cuckservatives" (Arciniega, 2017). In other words, the Alternative Right can be classified as an extreme right movement that claims to provide an "alternative" to classic conservatism. Without a clearly designated leading figure, a founding manifesto or a central point of reconvention, Alt-right represents a nebulous online grouping that is scattered across the many layers of the Internet, advancing its ideology on blogs, podcasts and online news-magazines none of which have the power to determine the direction of Alt-right entirely on their own (Marwick & Lewis, 2017). As this exact nebulousness has proved extremely beneficial for the development of the movement, the following chapters will further aim to examine whether the abstinence of the Alternative Right from a traditional organizational structure is a strategic or a coincidental development.

From #Gamergate into the White House

The group rose to relevance and had bolstered in numbers significantly in 2016, because of its involvement in the United States' presidential elections and suspected affiliation with Donald Trump's

election campaign (Neiwert, 2017). For the Alt-right, the highly politically incorrect presidential candidate represented an effective tool for splitting the Republican establishment and disrupting liberal harmony (Arciniega, 2017). Donald Trump's atypical and arbitrary approach to politics, and his candidacy against the embodiment of the liberal feminist establishment, Hillary Clinton, made him extremely attractive for the Alt-right (Arciniega, 2017). However, the Alternative Right did not endorse Trump's presidency because of its alignment with his politics; rather, the movement regarded the presidential candidate as a highly effective channel for shifting the limits of acceptable discourse, theoretically referred to as the "Overton window", to the extreme right in order to ultimately include the logic of eugenics into mainstream political discourse (Marwick & Lewis, 2017).

Following Donald Trump's victory, the far-right movement entered a union with more moderate figures on the conservative spectrum (Neiwert, 2017). Milo Yiannopolous and Joseph Watson, affiliated with Breitbart News Network and InfoWars respectively, aligned themselves with the Alt-right and resorted to the use of mocking and provocative language in attacking the political left and all those satisfying the label of a "cuckservative" (Amend & Hankes, 2018). Breitbart News Network and AltRight.com soon became the prime hub for the spread of pro-Trump propaganda (Marwick & Lewis, 2017). The publications placed on these websites downplayed many of the groupings nazi-esque elements, often attributing its racism to the dominance of the trolling culture within the online movement (Marwick & Lewis, 2017). In 2017, while still working at Breitbart, Bannon was appointed as the White House Chief Strategist in the office of U.S. President Donald Trump (Neiwert, 2017).

Alt-Internet: Alt-right's Parallel Virtual Universe

With the #Gamergate scandal fizzling out and under the umbrella of the Alt-right, the trolls, the angry gamers and the men's rights activists entered a union and became known as a movement recognizable by its self-referential, confusingly ironic, conspiratory and Nazi-inspired repertoire (Salazar, 2018). As prominent figures of the #Gamergate dispute rose to relevance, especially with Steve Bannon making it into Washington, a

quiet development had taken place: the male tribe formed under the hashtag #Gamergate had developed its own version of the Internet (Massanari, 2017). Now, the newly formed Alternative Right not only referred to a specific online culture of trolling, provocation and male tribalism, it also started to denote a collection of loosely connected blogs, news websites, subreddits, 4chan interest boards, forums and Twitter personalities (Massanari, 2017). Pre-existing websites like Breitbart and Altright.com now became a hub for right wing thought and expression. The online forum Reddit became host to countless sub-Reddits, like "Incels", "The Donald" and "Men Going Their Own Way", providing platforms to discuss a wide variety of niche topics (Sonnad & Squirrell, 2017). Within the span of a few months, the Alternative Right had established its own virtual empire (Marwick & Lewis, 2017).

One of the most remarkable sub-branches of Alt-right's parallel version of the Internet is the so-called "manosphere": an agglomeration of blogs and subfora dedicated to the discussion of male rights and all things masculine. On these websites the men are consolidated through their shared hatred for feminists. The prevalent narratives across the manosphere incorporate liberal tropes of oppression to propagate a perspective in which men, more specifically white cisgendered men, are oppressed at the hand of feminists (Marwick & Caplan, 2018).

The conspiratorial faction of the movement had also infested various blogs and forums, as well as mainstream websites such as YouTube. Without the barriers and standards set by traditional news media, the conspiracy thinkers created their own news shows and started publishing their own "documentaries", with the website InfoWars being one of the most prominent examples (Marwick & Lewis, 2017). The sentiment shared by the disseminated conspiracy theories expressed distrust in a powerful organization, the role and influence of which in an event or a given sphere had remained a secret. One of the biggest insecurities reoccurring across most of the conspiracies articulated a fear of white genocide or extinction of white culture and heritage(Marwick & Lewis, 2017; Neiwert, 2017)

Alongside the manosphere and conspiracy websites, the Alternative Right invested itself into establishing and further developing hyper-partisan online news networks such as Breitbart, InfoWars,

the Daily Caller, Truthfeed and Council of Conservative Citizens. These websites played an important role in the United States presidential elections, as their publications generated more engagement on websites like Facebook than the publications of mainstream news media (Marwick & Lewis, 2017).

In a matter of months, the previously disjointed group of men erected an interconnected network comprising their virtual "mancave" that has become a welcoming hub for many branches of far-right ideology (Marwick & Lewis, 2017). Bringing together individuals who share relatively radical opinions and providing them with space for knowledge dissemination and collective identity development (Marwick & Lewis, 2017). The social, participatory and deliberately disorganized nature of these networks poses a threat to the legitimacy and autonomy of mainstream media, making them vulnerable for manipulation (Marwick & Lewis, 2017). The instrumentalization of these blind spots of mainstream media and technological affordances of the Internet will represent the main object of study of the following chapter.

Chapter Two: Alternative Right: From 4chan to Charlottesville

On August 11 2017, a group of 250 young white males came together on the Nameless Field, a large patch of grass stretching out behind the Memorial Gymnasium on the campus of the University of Virginia. Under the command of their leader they formed a long column, lined up in pairs, each person holding a burning tiki torch. With instructions pouring in from multiple sides, the mass started marching and roaring slogans: "Our Blood, Our Soil", "You Will Not Replace Us", "Jews Will Not Replace Us" (Heim, 2017). With these hateful mantras echoing in the centuries old halls of the university, the marchers made their way to the hallowed Lawn where the counter-protesters had gathered around a Thomas Jefferson statue. In the dark, the torchbearers were confronted by a small group of University of Virginia students. Black counter-protesters were greeted with monkey noises (Heim, 2017). Soon the confrontation turned physical: punching, pulling, spitting. Multiple lit torches were thrown at the statue and the students before law enforcement got involved (Aston, Caron & Victor, 2017).

The following day, around 8 am, white supremacist activists filled the streets of Charlottesville to protest against the removal of a monument honoring Confederate hero, Robert E Lee. According to the voices of the far right, the removal of the statue was a metaphor for white genocide and the cultural and ethnic erasure of whites (Strickland, 2017). "You will not replace us!": chanting these words the protesters arrived at Emancipation park where the Confederate monument was located. Many were armed with firearms, bats and clubs; others proudly waved the Confederate flag and carried banners (Heim, 2017; Strickland, 2017). Both sides were throwing stones, ink-filled balloons, and bottles at each other in sporadic outbursts of violence (Heim, 2017; Strickland, 2017). Before the rally goers reached McIntire Park, the city of Charlottesville had ordered a state of emergency, shutting down the rally before it officially started (Heim, 2017; Strickland, 2017). On 4th street, several kilometers from the McIntire park, a 20-year-old white male James Alex Fields Jr. rammed his truck into the crowd, injuring 35 people and killing one. In the following days, the Charlottesville police department took in over 250 calls reporting assault from Saturday's rally (Heim, 2017). Despite the violent events, following the protests the Alt-right leader Richard Spencer informed the press on Monday that he would encourage his followers to return to Charlottesville to continue the protest against the removal of the statue, and denied any accountability of the Alternative Right for the violence of the past weekend (Heim, 2017).

The Gamergate scandal has significantly influenced the advancement of Alt-right as a movement by setting the conditions according to which the group would conduct the recruitment of its members and develop of its ideology. While the Alt-right represents only one of the many radical groups that have emerged in the global West in the recent years, its skillful utilization of online channels testifies to its uniqueness. Despite having emerged only four years ago, today the Alt-right is one of the biggest political movements in the United States. The fact that the Alt-right began with a self-identified disenfranchised group of gamers and turned into a political movement capable of mobilization demonstrates the relevance of online media to the democratic process and their ability to transform political landscape (Lyons, 2017). Nevertheless, while it is important to study the online

presence of the group, the Alt-right of today is no longer constituted by a group of angry 4chan users. The movement has long outgrown the virtual domains where its roots lie and its army of trolls represent it in protests and marches, armed with tiki torches and wrapped in confederate flags. While media attention and academic interest have been invested in the ideology of the group, the violent arm of the American Alternative Right has managed to remain under the radar. The following chapter will examine the Alt-right's journey from discursive violence toward collective physical violence. In doing so it will examine how the Alternative Right has managed to elevate itself from a geographically and ideologically scattered group of Internet users into a social movement that shares the same identity. This will be done through analyzing how the movement has adapted its online rhetoric to appeal to a larger audience, has actively fostered of a "us" versus "them" sentiment among its members and has been able to create a language of its own.

Storytelling Done Right: Implicit Hate and Intellectual Racism

Fostered by extreme political and ideological polarization in the United States following Barack Obama's presidency and the election of Donald Trump in 2016, today, the Alternative Right embodies a place of refuge for those who feel betrayed by mainstream politics and classical conservatism (Futrell & Simi, 2017). Staying true to its roots, the white supremacist movement has instrumentalized the Internet in the recruitment of its members and the dissemination of its ideas. The group has spun a strong online network that stretches over a multitude of websites. Some of these online platforms directly cater to the Alt-Right community: Altright.com and Breitbart.com are a good example of such networks. Furthermore, the self-identified members of the Alt-right are active on

forums such as Reddit and 4chan as well as mainstream social media such as Facebook and Instagram (Lyons, 2017).

The wide accessibility of these online platforms, and in many cases the possibility to remain anonymous has largely assisted the Alt-right in its growth and has helped it to de-marginalize its ideology by allowing it to introduce it to the mainstream. At the same time, the shift of white supremacy into the virtual dimension has imposed certain limitations and rules on the Alt-right's politics of hate. In their research on the oldest white supremacist website, Stormfront, Priscilla Marie Meddaugh and Jack Kay (2009) concluded that when right-wing groups spread their messages online, these messages tend to assume less blatant and aggressive formats when compared to traditional forms of hate texts. Ultimately, scholars argue that changed format of racism is influenced by the medium of online websites, and their open accessibility (Meddaugh & Kay, 2009). In other words, and in relation to our case study, the ability of anyone with internet connection to visit Alt-right websites to access, store and further disseminate the presented content forced the Alt-right to conduct its online politics with care. In practice, this meant that if the Alt-right did not want to scare away potential members with blatant racism, the movement had to rely on effective writing and pseudo-intellectual reasonings on racial hierarchy in order to expand its influence. In this sense, Meddaugh and Kay conclude, these websites represent a bridge between blatant hate and intellectual racism.

The following chapter will present a comparative analysis of comics created and disseminated under the White Aryan Resistance (WAR) between 1980 and 1990 and two articles published on a hyper-partisan news website AltRight.com. As the Altright.com website has been taken down in May 2018, the articles were retrieved from online internet archive "WayBackMachine" (Oppenheim, 2018).
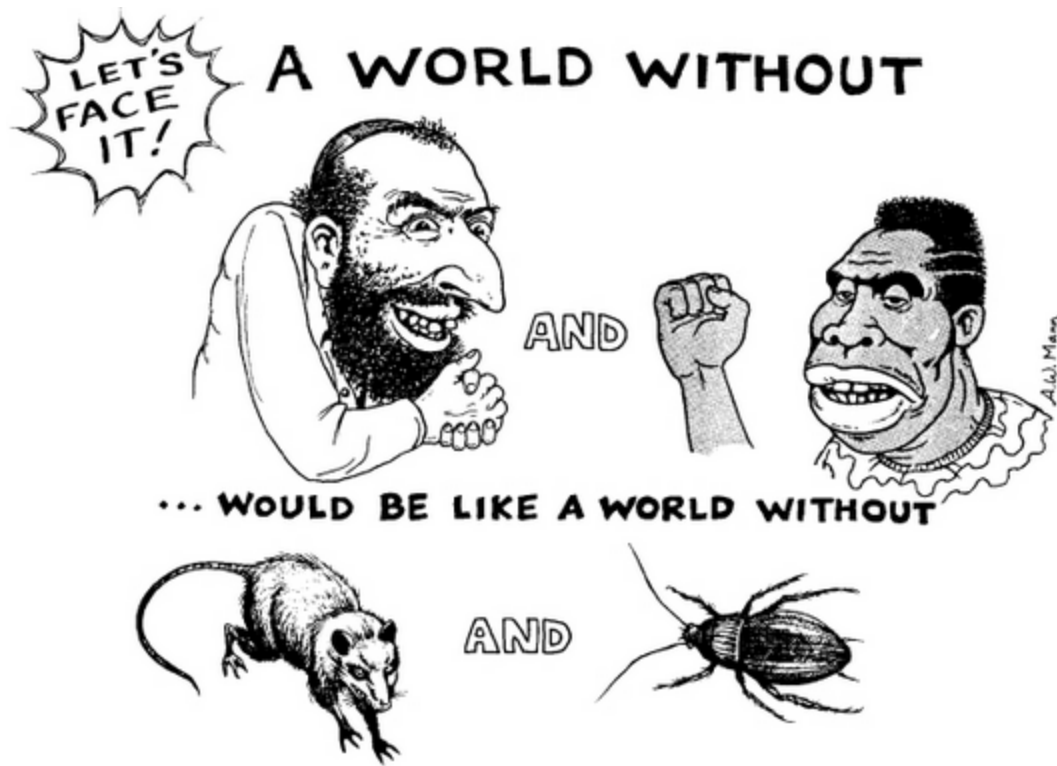
"WAR" comic (A. Wyatt Mann, n.d.)

Figure 1, presented above, depicts a comic drawn by Nick Bougas. Between 1980 and 1990 the cartoonist, better known under his pseudonym A.Wyatt Mann, produced numerous comics for the white supremacist movement the White Aryan Resistance (Bernstein, 2015). The comic asks its viewer to admit that "a world without [Jews] and [black people] would be like a world without [rats] and [cockroaches]". In doing so, the comic moves the viewer to fill in the blanks, thus making him complicit. In his portrayal of the Jewish and the black character, Bougas borrows elements from Nazi propaganda (Figure 2) and Thomas Rice's Jim Crow (Figure 3). This visible inspiration by imagery

Poster for a German Nazi Propaganda Film. The Eternal Jew (Holocaust Research Project, 2008)

The character of Jim Crow (Wikipedia, n.d.)

that is widely accepted as being racist, reveals that there was no intention to conceal the politics of the comic. Finally, by comparing Jews to rats and black people to cockroaches, the artist dehumanizes the "Other" that is embodied by the two ethnicities that the characters are representing.

Appendix A refers to an article published on Altright.com on November 7, 2017. The author of this piece, Gregory Hood, calls the president of the United States Donald Trump to "embrace his inner Andrew Jackson", the seventh president of the United States and well known for forcibly removing Native American populations from their homes and sending them to reservations in the American West. The article delineates an argument on Trump's Presidential Proclamation 9645, colloquially known as the third Muslim Ban (Hood, 2017). The travel ban addressed in the article restricts nationals from eight predominantly Muslim countries from entering the United States. The author specifically criticizes the attempts of the United States' Supreme Court and federal judges in preventing the presidential proclamation from taking effect, and advises Trump to follow in the footsteps of Andrew Jackson by restricting the power of the judi-

ciary branch:

> What Andrew Jackson said of the Bank of the United States must now be applied to the courts: "You are a den of vipers and thieves. I have determined to rout you out, and by the Eternal, I will rout you out!" (Hood, 2017)

Needless to say, this article is far from the overt racism present in A. Wyatt Mann's comic. However, without being unabashedly racist or hateful, the article manages to defend a very similar standpoint:

> The courts magically discovered "rights" to abortion and gay marriage. They declared *diversity* is a "compelling state interest" requiring government discrimination against whites. They threw out Proposition 187, declaring that illegal immigrants had a *right* to welfare from the very government whose laws they violated, and sealing the fate of the once great state of California. They destroyed American schools and mandated integration on reasoning that is now widely admitted to be faulty. [...] Now, we are told hostile foreigners have a right to come into the country [...] (Hood, 2017)

In the above paragraph, Hood echoes Alt-right's anti-immigrant, pro-life and homophobic outlooks. However, he justifies these views through providing concrete examples and referring to legal documents. He further argues that the provision of rights and privileges to immigrants, who by illegally entering the US violated the law, infringes upon the liberty of the law-abiding whites. The overall implication is thus that hostile foreigners do not "have a right to come into the country". However, instead of equaling the "hostile foreigners" to rats and cockroaches, Hood reaches out to legislation to support a standpoint similar to the one advocated by A. Wyatt Mann in his cartoon. At the same time, next to defending a political view, the article skillfully plays with aspects of Alt-right identity. The reference to Andrew Jackson plays a double role within the piece in that it not only helps the author to support his argument, but it also helps to bring to memory elements of American historical heritage, intending to evoke a certain sense of pride within

the reader. Finally, in the construction of his narrative and in support of his argument Hood mentions the names of John Locke and John Rawls:

> "Here, we see the rotten fruit of John Locke and John Rawls, after being dumbed down to serve an affirmative action culture. The state is conceived as a "neutral" arbitrator that assigns bundles of rights to every upright biped on the planet, without regard or to culture, capability, or, at this point, even citizenship. Never mind that John Locke and John Rawls themselves would not have agreed with it." (Hood, 2017)

The lack of proper explanation for the introduction of these philosophers into his piece testifies to Hood's presupposition of the familiarity of the reader with the mentioned thinkers. This assumption helps to construct a feeling of shared intellectuality, ensuring the readers that their views are not coming from a place of ignorance, but rather intellectual enlightenment.

The above strategy, defined as "implicit story telling" by Meddaugh and Kay (2009), not only allows the voices of Alt-right to uphold the aura of "intellectuality", but also promises that the presented interpretation of reality is rooted in logical reasoning. In researching Stormfront, one of the oldest Nazi websites, Meddaugh and Kay found support for the benefits implied by avoidance of directly inflammatory language, pointing out that "a tempered discourse that emphasizes pseudo-rational discussions of race [...] casts a wider net in attracting audiences" (2009, p.254). In his research of discursive power in far-right media on the Internet, Chris Atton (2006) reiterates Meddaugh and

Kay's observations, further suggesting that the implied "intellectual enlightenment" advocates that the websites used by these groups disclose hidden, alternate truths, and therefore can be identified as the "watchdogs" of democracy. Thus, where cartoons similar to those produced by A. Wyatt Mann would perhaps only appeal to a devoted racist, articles such as the one written by Gregory Hood are intellectually intriguing and attract individuals who otherwise would not have identified with the movement or would have been considered to fit its archetype (Lee & Leets, 2002) .

White Genocide : Interpretational Frames Within Alt-right Narratives

Ultimately, the efforts of Alt-right to make their rhetoric appealing to a wider audience are directed toward the creation of a shared collective identity. The implicit narrative construction yields space for persuasive, rich storytelling and has been found to have a more lasting effect on its receptors than imagery and texts that use aggressive and inflammatory language (Lee & Leets, 2002). In their account on white supremacist oppositional culture online, Josh Adams and Vincent Rosigno (2005, p. 760) suggest that collective identity formation is a necessary process that enables "a disparate group of individuals to voice grievances and pursue a collective goal under the guise of a unified empirical actor". It is argued that, through the creation of this shared identity, movements like the Alt-right can mobilize their members. Utilizing online fora, bulletin news boards and mainstream social media as a platform for community building and a means to connect previously disjointed individuals, the leaders of the movement design frames within

WAR comic (A. Wyatt Mann, n.d.)

their narrative building in such a way that the latter reflect "greater social values" while simultaneously addressing the pre-existing grievances and experiences of their existing members (Adams & Roscigno, 2005; Tarrow, 2011). In other words, the use of these frames of interpretation is necessary for the Alt-right to recruit sympathizers as well as maintain their support.

One of the reoccurring interpretational frames within the online rhetoric of Alt-right is the protection of white rights. The latter notion is deployed strategically by white supremacist organizations to incite a sense of urgency to act in the face of "white genocide". The idea is historically prevalent throughout the politics of white supremacist movements who have instrumentalized this notion to suggest that whites, and in the contemporary context especially white males, are an oppressed minority who are "deprived of exploring and celebrating their identity" (Lyons, 2017; Neiwert, 2017).

Figure 4 shows another comic drawn by A. Wyatt Mann between 1980 and 1990. The cartoon titled "We're nearing the end of the line" warns "White America" for the approaching "white extinction". The J&L (Jewish and Liberal) train is operated by Morris Dees, the co-founder of the Southern Poverty Law Center (SPLC). SPLC is a non-profit legal advocacy organization that has built a reputation through its successful legal cases against hate groups like the Klu Klux Klan and for its efforts to set up educative tolerance programs. The individual wagons of the train mention some of the individuals or entities who are heading toward "extinction". Among those individuals the founder of White Aryan Resistance and a former Klansman, Tom Metzger.

A similar thought is reiterated in Appendix B. that refers to an article published on Altright.com on 28th of October in 2017. The writer of the column "Adios, California" and a guest contributor Melinda Santa Cruz voices her concern for the drastic demographic shifts that she has been witnessing in the state of California since the 1970's. In the following paragraph the author describes her per-

sonal experience with illegal immigrants from Mexico:

> My mother did hire an illegal couple to work on the ranch. [...] We knew they were illegal and it royally pissed me off because they were taking advantage of our good will without having regard for our laws. But they weren't bad people overall. In the 1990s [...] the days of having peasant farmers crossing the border for a better life were over. We now have hardcore criminals committing murder and rape. [...] My grandfather was the only white person left in the town where he had lived for sixty years. (Santa Cruz, 2017)

Here, Santa Cruz describes the sudden demographic shift that she had witnessed, noting that also the *type* of immigrants who crossed the border has changed from farmers to "hardcore criminals committing murder and rape" (Santa Cruz, 2017).

To visualize her concerns, she presents an experience of her own:

> My daughter was preyed upon by Mexican gang-bangers who thought she was one of them. Her father was Spanish, not Mexican. She was called a "disgrace to her race" and told to stop "acting white". "What are you talking about?" she asked. "I am white". Her European ancestry made no difference to them (Santa Cruz, 2017).

The passage reveals that the writer had, until the occurrence of the described incident, assumed that European ancestry could fulfill a protective role, in that it would shield her daughter from interaction with the members of minority communities. Instead, she found that her daughter's performance of "white" identity was suppressed by "Mexican gang-bangers". Having justified her source of frustration with Mexican immigrants through telling a personal anecdote, Santa Cruz moves on to presenting her interpretational frame:

> I transferred her to another high school only to find that Mexican students were being bussed into that white neighborhood. I had to drive her there every morning because there was no bus for her, but there were two for them (Santa Cruz, 2017).

Through mentioning the necessity to drive her daughter to school daily, while Mexican students were offered school busses the author provides a practical example of discrimination against whites. Where about four decades ago Mexican immigrants were "poor peasants looking for a better life", today they are dangerous "gangbangers" who enjoy privileges that are not available for white people, who are aboard on the train heading toward "white extinction". Aside from being an example of the notion of white victimization, the passage above also demonstrates how this frame helps to establish an us versus them dichotomy. While the concerns voiced for the future of white people address a shared grievance and help the manifestation of a collective identity, the antagonizing portrayal of Mexican immigrants serves to define the "Other" within the narration.

Blue Pill, Red Pill : Speaking the Language of the Alt-right

Another affordance of digital platforms is that they can host numerous amounts of people at the same time, coordinating disjointed individuals into a united entity by allowing them to enter a democratic discourse. Public message boards like 4chan, 8chan and Reddit are well suited to accommodate discussions and represent digital spaces where the Alt-right's online culture came to develop (Marwick & Lewis, 2017). An interesting pattern that has unfolded on these websites is the creation of an Alt-right jargon. The coming into being of this exclusive glossary has stimulated the manifestation of a shared Alt-right identity as well as helping to exclude those who are not familiar with its lexicon (Meddaugh & Kay, 2009). Beyond representing an "Othering" mechanism, this language reveals the issues that are deemed important by the movement to the extent that they necessitate the creation of a new word (Adams & Roscigno, 2005; Meddaugh & Kay, 2009).

News websites Quartz, Nikhil Sonnad and Tim Squirrell (2017), recently published an article containing a list of vocabulary common to the Alt-right lexicon. To produce this list, the authors studied various Reddit subforums (outlined below) and

identified the active contributors of these so-called "boards" who have demonstrated an affiliation with the Alt-right.

1. Conspiracy, where conspiracy theories are traded with the utmost seriousness

2. Incels, or "involuntary celibacy", where men blame society for their romantic failings

3. KotakuInAction, home of the somehow-still-active #Gamergate movement, and where video games are discussed

4. MGTOW, or "Men Going Their Own Way,"f where men decry feminism

5. CringeAnarchy, where users cringe at social justice warrior artifacts around the internet

6. The_Donald, the biggest Alt-right hangout, where serious Donald Trump supporters congregate (Sonnad & Squirrel, 2017)

Words like blue pill, red pill, cuckservative, femoid and social justice warriors (SJW) were found to be among the most popular. The concept of the blue and the red pill is inspired by the 1999 science fiction movie The Matrix (Sonnad & Squirrell, 2017). In the movie, the protagonist Neo is asked to choose between a red or a blue pill. The red pill stands for a painful awakening and would make Neo see the world as it really is; taking the blue pill would allow Neo to remain in the state of comfortable ignorance (Sonnad & Squirrell, 2017). Within the Alt-right framework, being "redpilled" means to discover a hidden truth, coming to see the world as it is and realizing that political correctness is a tool used by SJW's to suppress white culture and undermine the Western civilization. The terms "cuckservative" or "cuck" are used to address and emasculate mainstream conservatives. Originating on Reddit's Incels board, the word "femoid" combines the words "female", "humanoid", which itself is a portmanteau of "human" and "android", and allows the self-identified "incels" to refer to women as sub-human creatures. Beyond identifying the issues that are deemed important by the Alt-right, these words also reveal Alt-right sentiments surrounding these same issues (Sonnad & Squirrel, 2017). For example, being red pilled doesn't only mean discovering the truth; it also implies that the truth was hidden away either by the establishment, SJW's,

feminists or liberals. In consequence, the demonization of the so-called social justice warriors, feminists and liberals is normalized since these are the same people who have wronged the Alt-right. The word "femoid" is equally disturbing, and the integration of this word within the rhetoric of the Alt-right reveals how permissive the environment is toward derogatory stances against the female gender.

Studying these previously mentioned subreddits is also valuable as they reveal that the sympathizers of Alt-right ideology tend to self-select into different sub-branches of the movement. What makes this phenomenon particularly interesting, is that the shared cultural and social background, especially as found on specific subreddits, allows for a comfortable exchange of ideas, increasing the mutual agreement on issues and eventually shifting the average opinion in the group towards an ideological extreme (McCauley & Moskalenko, 2008). Radicalization experts Clark McCauley and Sophie Moskalenko (2008) denote this development as radicalization in like-minded groups, noting that the extreme antagonization of an "Other" within this social setting and its relevance for the vitality of the rhetoric is emblematic for a mass radicalization mechanism that is fueled by hate against an out-group. Collective identification within the in-group can drive its members to assume extremely hostile positions even, or especially in, the absence of personal grievances caused by the opposition (McCauley & Moskalenko, 2008).

From Aggressive Discourse Toward Physical Violence

The ability of the Alternative Right to maneuver across a large variety of media channels and weaponize discourse to propagate a worldview that denotes and antagonizes a definite set of "enemies" can be classified as an act violence. In his 1969 thesis Johan Galtung, a Norwegian sociologist, and one of the most prominent names in peace and conflict studies, articulates a definition of violence that does not necessitate for an agent to be hit or hurt in order to assert that they have been violated. Rather, he suggests that the existing threat of mental or physical violence with the potential to constrain human action qualifies an act as violence. He further asserts that violence can more precisely be identified either as per-

sonal or structural depending on the presence of an agent who commits it. As personal or indirect violence he defines those acts where an agent is involved who commits the violence. The violence that is present within the existing structure, i.e. the society and the institutions, and is manifested as "unequal power and consequently as unequal life chances" is denoted as structural violence (Galtung, 1969, p.171). Concerning the relationship between these types of violence, Galtung argues that personal violence shows when this structure starts disintegrating as "those who benefit from structural violence [...] will try to preserve the status quo so well geared to protect their interest"(Galtung, 1969, p.179). Thus, when the existing system is threatened, through examining the behavior of social groups and their commitment to preserve this system one can identify the hierarchical belonging of an individual (Galtung, 1969). In the case of the Alternative Right, and specifically taking into account the values of the movement, it is apparent that the grievances voiced within its narratives often relate to social changes brought about by globalization. For example, the article analyzed earlier (Appendix 2) reveals that the writer is bothered by the strong demographic shift in the state of California. The Altright.com contributor specifically voices her discontent with the provision of school busses for Mexican students, while she is forced to drive her daughter to school, as such service is not provided to white students. The discontent articulated in this piece testifies to the belief held by Alt-right sympathizers that the existing structure is threatened. Having asserted her frustration with the absence of a school bus, the author finishes her paragraph saying this: "And the globalists wonder why we voted for Trump?" (Santa Cruz, 2017). In the light of Galtung's theory, the vote for Donald Trump can be seen as an effort to maintain the previously beneficial structure. It is important to mention however, that within the Alt-right's narration, white people are not depicted as becoming less privileged, but are said to be actively oppressed, thus representing the segment of society that is suffering under structural violence. Galtung further suggests that in attempts to restore social justice, the benefits of short-term personal violence are deemed to outweigh the costs of persistent structural violence. However, the acceptance of personal violence as a means to attain social justice often leads to the acceptance

of physical violence, and Galtung argues that violent revolutions "testify to this" (Galtung, 1969, p. 184). The ability of the Alternative Right to establish a powerful and dominant frame of interpretation to propagate the white nationalist agenda has attracted countless allies. In consequence, despite the loose organizational structure of the movement the collective is linguistically distinguished as the "Alternative Right". The public acknowledgement of Alt-right as a designated collective, unintentionally mobilizes the movement and lends it with the competence for collective action. The skillful framing of the grievances experienced by some of America's whites as systematic oppression and structural violence incites immediate action, thus condoning personal and consequentially physical violence. The violent aftermath of the Unite the Right protests in Charlottesville testifies to this and provides a glimpse into the inevitably violent future of the Alternative Right.

Chapter Three: Learning the Language of the Media: Alt-right gone Mainstream

As the previous chapter demonstrates, the Alternative Right's repertoire of collective action is focused on the advancement of a specific symbolic frame, namely the implied threat of multiculturalism on white culture and the white race. We have seen that the movement strongly relies on the Internet to disseminate its message and normalize its frame of interpretation (Tarrow, 2011). Not only is the Internet unique in that it is one of the most inclusive, accessible and affordable forums for the exchange of information, but the medium itself fosters a feeling of a community and interconnectedness (Adams & Roscigno, 2005; Meddaugh & Kay, 2009). Further, compared to printed press and television, the Internet represents a space that experiences the least political and media regulation, thus proving to serve as an efficient tool for the spread of ideological and political propaganda (Marwick & Lewis, 2017). Chapter two also illustrated how the Internet has been instrumentalized by the Alternative Right to construct a shared collective identity, recruit members and maintain their support, demonstrating the revolutionary role of this medium in allowing collective formation to take place in private spheres rather than in public. At

the same time, while the Internet has played a crucial role in the development of the Alternative Right, the movement has not limited itself to this medium; in fact, specifically looking at the growth of the Alt-right since the last American presidential elections, it is apparent that mainstream news media outlets have been decisive in the success of the group. In his analysis on the construction of social protest the sociologist William Anthony Gamson (1975, p.85) refers to movement activists as "media junkies" who "monitor the public discourse" to "construct meaning on issues they care about". Specifically, he says, these "media junkies" are skillful in bridging public discourse with the personal experiences of people, thereby creating symbolic frames of interpretation that compel collective action (Gamson, 1975). The following chapter will look at the ways in which the Alternative Right has manipulated and instrumentalized the affordances of mainstream media to advance its goals and normalize its way of thinking.

Bad Hombres :   How Donald Trump Shifted the Overton Window

Within the mainstream political discourse, the Alternative Right is often associated with the United States' President Donald Trump, most commonly because of the movement's alleged endorsement of Trump. However, as is mentioned in the first chapter of this thesis, the group did not favor Trump because of his political standing, as is widely believed, but because his controversial statements lead to the widening of the range of thoughts tolerated within the public discourse, also referred to as the shifting of the Overton window. This notion refers to an observation made by Joseph Overton, the senior Vice President of the Mackinac Center for Public Policy. Overton noted that within certain public policy areas only a limited amount of ideas are considered "politically acceptable" (Lehman, n.d.). In the political context, the viability of an idea is therefore almost entirely dependent on the width of the Overton window and not as much on the preference of the politician. In this sense, the Overton window shapes the preferences of the politicians in that it provides a guideline for what the politicians believe "they can support and still win reelection" (Lehman, n.d.). Consequently, the shifting of the window takes place not when the preferences of the politicians change, but when the ideas

and preferences change in the society that elects these same politicians. In other words, it is the public's conception of the window that dictates that of the politicians? and not the other way around, as is commonly assumed. In practice, this means that there exists a certain margin of acceptable ideas within the public discourse, ideas situated outside of that margin are considered radical. Exposing people to extreme ideas positioned far from the borders of the constructed window, regardless if they accept it, leads them to consider relatively less radical ideas as more acceptable, thus pushing the window in a slightly radical direction. In general then, the shifting of the Overton window is not done to "normalize" extreme ideas, but rather to gradually move it towards those extremes by stimulating the political imagination of the electorate.

This means that when Donald Trump warned the American public against the "bad hombres" and rapists entering the country via Mexico, he did not actually "normalize" this type of antagonizing rhetoric, rather these comments made his previous statements seem less radical (Agren, 2017). It is not hard to imagine that the Alternative Right has enthusiastically embraced this concept. Some of the figures situated on the forefront of the Alternative Right, like Paul Nehlen, even went as far as to claim the shifting of the Overton window as one of the primary goals of the movement (Holt, 2017). The shifting of the Overton window largely facilitated the transition of Alternative right's white supremacist agenda from the virtual margins into the mainstream media.

Don't Feed the Trolls : Pizzagate and Lulz

Before looking at how the Alternative Right has made the shift from alternative into the mainstream media, it is important to once again consider the aspects of the movements digital background. To refresh our memory on the origins of the Alternative Right: the movement started gaining momentum just as the #Gamergate scandal started unfolding on websites like 4chan and Reddit. Coming from this background, the Alternative Right is founded on the tradition of online trolling and meme culture that emblematic of 4chan and Reddit. In the previous chapter, we have observed that the appropriation of the digital space by right wing groups has prompted a shift in the tone of modern hate speech. In a similar manner, the affor-

dances of 4chan as a web forum have defined what we today know as trolling. The home page of 4chan indicates that the website is "a simple image-based bulletin board" and that its users are not required to register to contribute: "feel free to click on a board below that interests you and jump right in!" ("What is 4chan?", n.d.).

Trolling refers to a repertoire of unsolicited deliberate online behavior that is directed to generate an emotional reaction. In her book "This is Why We Can't Have Nice Things" Whitney Phillips (2016, p. 10) describes trolls as tricksters, who "take perverse joy in ruining a complete stranger's day". Here, Philips points out one of the driving dynamics behind trolling: the asymmetrical relationship between the troll and the trolled. Within this dynamic, the trolls deny accountability for their behavior through their mantra "nothing should be taken seriously", thus claiming that even when abusive or inflammatory language is part of their repertoire, it does not mean that the troll is serious or has to be serious about what he or she says (Phillips, 2016). When the victim of trolling attempts to call out the troll on their behavior, they are trolled even worse. Hence, if the target of trolling gets angry and addresses the behavior of the troll, he is "feeding" the troll and satisfying his chief purpose: generating anguish. This asymmetry between the troll and the target of trolling and the refusal of trolls to provide the treatment they demand from others show how trolling is the ultimate assertion of privilege (Phillips, 2016). The principles of irony, amusement and anonymity that 4chan is founded on, often accompany the practice of trolling and serve to alleviate the burden of accountability off of the troll. Further, an online aphorism, Poe's Law acts as an enabling mechanism for some of trolling repertoires: the "law" states that without a clear indication of the authors intent, such as a "winking smiley", something that was intended to serve as a parody of an extreme view will not be recognized as such, or even, will be seen as the extreme expression it is meant to mock (Aikin, 2013). Poe's Law goes on to show how virtual content, the intention of which is unclear, not only leads to the polarization of discourse, but also leads the consumers of virtual content to question their media literacy, fostering feelings of alienation and distrust among the users of online platforms (Aikin, 2013). A practical example of instrumentalized trolling and the weaponization of Poe's Law at the hands of the Al-ternative Right is the so-called "Pizzagate" incident.

During the 2016 United States' presidential elections, the conspiracy circles on 4chan and Reddit were leading months long discussions about the alleged involvement of the democratic candidate, Hillary Clinton, in satanic rituals as well as a child sex ring (Marwick & Lewis, 2017). The conspiracies mostly enjoyed popularity among the users of these platforms, only occasionally appearing in the form of news on small scale websites that were seeking to generate advertisement revenues with sensationalist headlines (Marwick & Lewis, 2017). Throughout the election cycle, the conspiracies made their way onto mainstream social media websites like Facebook and Twitter, where they gained thousands of likes, shares and retweets. The turning point within the Clinton conspiracy cycle was the release of Hillary Clinton's private hacked e-mails on WikiLeaks in October 2016, only a month before the presidential elections (Marwick & Lewis, 2017). Following this release, the e-mails became the main point of discussion on subreddits like The_Donald, and 4chan. In an effort to discover information that could help Donald Trump's presidential campaign, the users of these platforms collectively went through all of Clinton's e-mails (Robb, 2017). An email between Hillary's campaign chair, John Podesta, and the owner of a Washington DC pizza restaurant captured the attention of some 4chan users. The e-mail discussed the details of a fundraiser organized by Clinton's team that was going to take place at Comet Ping Pong in DC. The e-mail raised speculations: the phrase "cheese pizza" had previously appeared on 4chan as a code word for child pornography, and some users suggested that food related terms in the leaked e-mails also had double meanings. An anonymous user speculated that "pizza" is a secret word for child sex and that Comet Ping Pong served as the headquarter for the alleged child sex ring, and this interpretation was soon popularized (Aisch, Huang & Kang, 2016) This theory, now called Pizzagate, was also picked up by an Alternative Right darling, and conspiracy thinker, Alex Jones, who discussed the conspiracy on his popular show InfoWars. In the following months Comet Ping Pong became the target of threats and its employees victims of violence. On 4th of December a gunman entered the restaurant, claiming that he came to conduct his own investigation (Robb, 2017). Following his arrest, the victim of Poe's law denied having ties with the Alt-

right, but confirmed that he has studied articles addressing the conspiracy and listened to Alex Jones, who played an important role in promoting the conspiracy. The incident received worldwide coverage from mainstream news media, drawing further attention to the Pizzagate conspiracy (Marwick & Lewis, 2017). By putting the Alternative Right in global limelight, the media fed the troll.

The most powerful aspect of the Pizzagate stunt is perhaps the demonstrated ability of the Alt-right to skillfully play with the definition of a conspiracy and blatant "fake news". A conspiracy often implies the devotion and sincere trust of the conspiracy thinker in the theory, whereas "fake news" is something that is disseminated with the intent of misinforming for personal advantage (Marwick & Lewis, 2017). Pizzagate was able to play with best aspects of both phenomena. Skewed facts and misinformation developed into a conspiracy theory in collaborative environments of 4chan and Reddit, and a sincere believer of this conspiracy was moved to pick up a gun and undertake individual action (Robb, 2017). The media coverage of the incident attracted more attention to the theory and the presumably dismissive attitude of the media towards the facts of Pizzagate reinforced the already marginalized feeling amongst the conspiracy thinkers (Marwick & Lewis, 2017). The Pizzagate incident is not only representative of the ability of Alternative Right to manipulate the media and spread misinformation, but it also shows how easily misinformation can spill over into real-life violence. Perhaps the most frightening lesson taught by Pizzagate is that in the aftermath of potential real-life violence we will not be able to assign responsibility to the actual perpetrator of the crime, because the troll will remain anonymous, and they probably "didn't mean it anyway".

Weapons of Mass Distraction : Framing and Media Logic

In the last decades, the media have acquired a pivotal role Western democracies' political processes, taking up central roles during election campaigns, political image building and public democracy (Mazolleni, 2008). In his account on the growing influence of media and their embeddedness in the political sphere, the communications scholar Gianpietro Mazolleni (2008) conceptualizes this development as the "mediatization of politics". He notes that in modern democracies, the news media increasingly represent the most reliable source of information and play an important role in establishing the conceptual and cultural frameworks according to which the public sentiment is formed and political reality is framed (Mazolleni, 2008). In more permissive environments where media are granted autonomy, they often hold the power to drive the public debate and influence the political agenda through selective coverage, use of different cultural frames and the attribution of different levels of salience to different issues. At the same time, today modern media are governed much alike commercial businesses, existing within capitalist structures with the ultimate goal to sustain themselves(Strömback & Esser, 2009). With their functions defined by commercial interest and in societies where news is free and only one click away, the media ultimately seek to remain commercially relevant and are thus mainly invested in capturing the sparse collective attention of the public. In addressing the problem of attention in digital age, the scholar Yves Citton (2017) establishes an "ecology" of attention. He outlines a few principles guiding this concept, the first three of which are relevant to the thesis at hand. As the first principle Citton states that attention must be conceived of as a finite resource. He then claims that attention can never be distributed equally to all occurring phenomena and that simultaneously occurring issuesare therefore in competition to receive attention. Finally, he states that one's attention usually is drawn more towards those phenomena that one is already familiar with (Citton, 2017). Following this logic, in attempts to capture the attention of the public, the media must consider the guiding principles of "attention" in news production and coverage. For example, by perceiving attention as a finite resource, the media translate important information into intriguing headlines to capture attention. To prevent the loss of this newly gained attention to a different phenomenon, the media resort to storytelling techniques such as simplification, polarization and stereotypification, thus counting on the already existing familiarity of the news consumer with one of these techniques (Strömback & Esser, 2009). The media further adhere to concrete and disjointed frames, that only require the attention of the viewer in the moment they are consuming the information as opposed to thematic and complex frames that demand uninterrupted at-

tention (Iyengar, 1991). To see how the new age of commercially driven news media translates into practice and how media indirectly tolerate and incite violence, let us look at how the American media covered the 2015 Charleston church shooting.

Alt-white : Towards a Race War

In June of 2015 a 21-year-old male, later discovered to be Dylann Roof, entered the doors of the Emanuel African Methodist Episcopal Church in Charleston, South Carolina and killed nine people during a prayer service (Workman & Kannapell, 2015). The online manifesto published by Roof prior to the shooting revealed that he was driven by a desire to ignite a race war and thanked the Alt-right affiliated website "Council of Conservative Citizens" for his "racial awakening" (Roof, n.d.). Warning the readers against the dangers of multiculturalism and calling for a conception of race as something "real" and not constructed, Roof's manifesto painfully reiterated many of Alt-right's core values (Roof, n.d.). The preservation of white culture and heritage, as the core of Alt-right's values, represented the essence of Roof's manifesto. Following the massacre, the Alt-right's unofficial leader Richard Spencer admitted that as much as the movement would like to, it could not write off Dylann Roof as a "madman", as his writing revealed capability for critical thinking and understanding of implications of race on the future of white America (Brown, 2017).

In the days following the massacre, influential mainstream news media like the Washington Post and The New York times referred to Dylann Roof to as the "Charleston Church Shooter", "Racist Church Shooter" or simply as "Dylann Roof" (Zapotoski, 2017; Blinder & Sack, 2016). In spite of the evidence of Roof representing a radicalized face of the Alternative Right, with the majority of news media outlets adhering to the widely recognized frame of a white lone gunman, only a few media outlets embraced Roof as a radicalized race ideologist. Guided by commercial interest, mainstream news media sources refrained from introducing new frames and did not make an effort to connect the Charleston church massacre with the rise of the popularity of Alt-right, thus alleviating the responsibility off the movement and presenting a version of this event as an individual occurrence disjointed from a cultural or social background.

The ability of the Alternative Right to identify and exploit the principles of traditional media discourse allows it to navigate through the media landscape as a legitimate participant and skillfully weaponize notions like the Overton window and Poe's law. At the same time, as the white supremacist movement originated within the domains of 4chan and Reddit and thus was born into an aura of irony and humor it is often recognized as a legitimate representative of the said subcultures. In reality, Alt-right has instrumentalized its familiarity with elements of these online cultures to subsequently repurpose practices such as trolling and meme sharing to advance its radical politics under the ploy of comedy and satire. As a result, the racism and bigotry of the Alternative Right have not gone unnoticed, but the nebulousness of the movement does not allow for the reconciliation of its many identities as racists, trolls, young, sharp-tongued men, an online mob that hides behind anonymous nicknames, Trump supporters, church shooters and livid misogynists. In the midst of an attention crisis and in the age of distraction, traditional media struggle to fulfill their task as watchdogs and retain commercial appeal. In the meantime, as the Alternative Right professes a "white sharia", the United States welcomes a new culture of violence, and no one is held responsible.

Conclusion

Public expression of racist thought is a universally rejected practice. Yet, it is through engagement in such practices that the Alternative Right has secured a spot on the edges of the mainstream. Despite the lack of a traditional organizational structure, the Alt-right represents a movement with a clear message and a set goal: white people are superior, they are oppressed, and the world must know this truth. The abstinence from hierarchical patterns is intentional, as it makes the Alt-right more difficult to trace for traditional media and the opposition. At the same time this avoidance of order facilitates intragroup communication and allows Alt-right to manifest itself as an agile, flexible and a highly networked organization.

By tracing the origin of the movement and assessing the events that have led to its birth, this paper has shown that the Alternative Right derives its many faces from the cultural surroundings it en-

ters or has previously occupied. The acceptance of aggressive speech as a means to an end, and violent resistance against any degree of infringement upon previously undisputed entitlement, has accompanied the Alt-right from its crib that was #Gamergate to the graves of the victims of its violence. Scattered across countless websites and blogs and geographically disjointed, the Alternative Right is united through a distinct tradition of discursive and physical violence. Adapting itself to the modern context and with regard to the affordances and limitations on modern media channels, white supremacy under the moniker Alt-right has made a shift from blatantly dehumanizing and slur-infused messages, to ones appealing to cultural nostalgia, appearing intellectually supported and bathing in an aura of semi-sophistication. Alt-right has created a mirror version of the Internet and has developed its own language, both meant to exclude the "Other" and construct an "us". The perceived contention of privilege and entitlement of white men brought about by socio-demographic shifts is translated as a threat to the future of white identity and is seen as an exercising of structural violence against whites. Consequently, the strongly propagated frame of marginalized whites serves to justify the violence caused at the hands of the movement. Further, the movement showcases exceptional familiarity, not only with the conventions of the Internet, but also traditional media discourse. Alt-right's literacy in these fields allows it to manipulate the media landscape into introducing it into the mainstream and allows it to build a public image that is softened by confusion and euphemistic language. Although the Alt-right's rhetoric does not explicitly call for physical violence, the place of mistrust and grievance from where the voice of the group originates serves as a justification for the resulting atrocities, which are not appropriately problematized. The Alternative Right's narratives emphasize the need to act urgently: the violent aftermath of the protests in Charlottesville and the deaths brought about by the hands of Dylann Roof are evidence of their effectiveness.

The scope of this thesis has allowed for the mapping of Alt-right's development from angry gamers to White House Chief Strategists and radicalized gunmen, the exploration of the movement's interpretational frames, as well as its success at legitimizing itself through the weaponization of the blind spots of traditional media. The issue of the gender and race dimensions represent two themes that have prevailed throughout this thesis but have not been incorporated into its main structure due to practical limitations. A greater focus on these two aspects would allow the topic of the Alt-right to render itself to the field of gender and critical race studies, and compels the assessment of the role of traditional ideas of masculinity and the implications of race in the development and ideology of the movement. Today, the number of people killed by the Alternative Right remains at 43, but with every red pill that is distributed, the number of victims rises. The time has come to tell America's Lost Boys that these violent delights have violent ends.

References

Adams, J., & Roscigno, V. T. (2005). White supremacists , oppositional culture and the World Wide Web. Social Forces, 84(2), 759?778. https://doi.org/http://www.jstor.org/stable/3598477

Agren, D. (2017, February 2). 'Bad hombres': Reports claim Trump spoke of sending troops to Mexico. The Guardian. Retrieved from http://www.theguardian.com

Aikin, S. F. (2013). Poe's Law, group polarization, and argumentative failure in religious and political discourse. Social Semiotics, 23(3), 301?317. https://doi.org/10.1080/10350330.2012.719728

Aisch, G., Huang, J., Kang, C. (2016, December 10). Dissecting the #Pizzagate conspiracy theories. The New York Times. Retrieved from http://www.thenewyorktimes.com

Amend, A., & Hankes, K. (2018, February 05). The Alt-Right is killing people. Retrieved May 27, 2018, from https://www.splcenter.org/20180205/alt-right-killing-people

Arciniega, L. C. (2017). Diversity and inclusion and the rise of the Alt- Right. Anthropology News, 58(1), 175?179. https://doi.org/10.1111/AN.332

Astor, M., Caron. C., Victor, D. (2017, August 13). A guide to the Charlottesville aftermath. The New York Times. Retrieved from http://www.nytimes.com

A.Wyatt Mann. (n.d.). Retrieved from http://knowyourmeme.com/memes/happy-merchant

A. Wyatt Mann (n.d.). Retrieved from https://forum.facepunch.com/f/videos/bqxyl/pol-raids-a-fox-news-chatroom-of-the-Baltimore-riots-as-the-fox-reporters-read-the-comments-on-ai1/

Blinder, A., Sack, K. (2016, December 15). Dylann Roof found guilty in Charleston church massacre. The New York Times. Retrieved from http://www.thenewyorktimes.com

Brown, J. (2017, January 12). Dylann Roof, the radicalization of the Alt-Right, and ritualized racial violence. The Martin Marty Center for the Public Understanding of Religion. Retrieved from http://www.divinity.uchicago.edu

Citton, Y. (2017). Ecology of Attention. Polity Press.

Der Ewige Jude (n.d.) Retrieved from www.holocaustresearchproject.org/holop/derewigejude

Futrell , R., & Simi, P. (2017). The [un]urprising Alt-Right. Contexts, 16(2), 76?76. https://doi.org/10.1177/1536504217714269

Galtung, J. (1969). Violence, peace, and peace research. Journal of Peace Research, 6(3), 167?191. https://doi.org/http://www.jstor.org/stable/422690

Gamson, W. A. (1975). Strategy of Social Protest. Social movements and culture, 4, 85-106

Goldstein. J. (2016, November 20). Alt-Right gathering exults in Trump election with nazi-era salute. The New York Times. Retrieved from http://www.nytimes.com

Heim, J. (2017, August 14). Recounting a day of rage, hate, violence, and death. The Washington Post. Retrieved from http://www.washingtonpost.com

Holt, J. (2017, December 18). Paul Nehlen is channeling the Alt-Right to shift the 'Overton Window'. Right Wing Watch. Retrieved from http://www.rightwingwatch.org

Hood, G. (2017). By the eternal. AltRight. Retrieved from https://web.archive.org/web/20180501012853/https://altright.com/

Iyengar, S. (1991). Is anyone responsible? How television frames political issues. Chicago, IL: University of Chicago Press

Jim Crow (n.d.) Retrieved from https://en.wikipedia.org/wiki/File:Jimcrow.jpg

Lee, E., & Leets, L. (2002). Persuasive Storytelling. American Behavioral Scientist, 45(6), 927?957. https://doi.org/10.1177/0002764202045006003

Lehman, J.(n.d.). A brief explanation of the Overton window. Retrieved from http://www.mackinac.org/overtonwindow#top

Lyons, M. (2017). CTRL-ALT-DELETE: The origins and ideology of the Alternative Right, 1?21. Retrieved from http://www.politicalresearch.org/wp-content/uploads/2017/01Lyons_CtrlAltDelete_PRINT.pdf

Marwick, A. E., & Caplan, R. (2018). Drinking male tears: Language, the manosphere, and networked harassment. Feminist Media Studies, 0777, 1?17. https://doi.org/10.1080/14680777.2018.1450568

Marwick, A., & Lewis, R. (2017). Media Manipulation and Disinformation Online. Data & Society Research Institute, 1?104.

Massanari, A. (2017). #Gamergate and the fappening: How Reddit's algorithm, governance, and culture support toxic technocultures. New Media and Society, 19(3), 329?346. https://doi.org/10.1177/1461444815608807

Mazzoleni, G. (2008). Mediatization of politics, The International Encyclopedia of Communication, 7, 3052-3055

McCauley, C., & Moskalenko, S. (2008). Mechanisms of political radicalization: Pathways toward terrorism. Terrorism and Political Violence, 20(3), 415?433. https://doi.org/10.1080/09546550802073367

Meddaugh, P. M., & Kay, J. (2009). Hate speech or 'reasonable racism'? The Other in Stormfront. Journal of Mass Media Ethics, 24(4), 251?268. https://doi.org/10.1080/08900520903320936

Neiwert, D. (2017). Birth of the Alt Right. The Public Eye, 1, 4?10. Retrieved from http://www.politicalresearch.org/2017/03/22/birth-of-the-alt-right/#sthash.0QQKjCnV.dpbs

Oppenheim, M. (2018, April 15). Facebook shuts down white supremacist Richard Spencer's pages. Independent. Retrieved from http://www.independent.co.uk

Perreault, G. P., & Vos, T. P. (2018). The GamerGate controversy and journalistic paradigm maintenance. Journalism, 19(4), 553?569. https://doi.org/10.1177/1464884916670932

Phillips, W. (2016). This is why we cant have nice things: Mapping the relationship between online trolling and mainstream culture. Cambridge, MA: The MIT Press.

Robb, A. (2017, November 16). Anatomy of a fake

news scandal. Rolling Stone. Retrieved from http://www.rollingstone.com

Roof, D. (n.d.) Dylann Roof's manifesto. The New York Times. Retrieved from http://www.thenewyorktimes.com

Santa Cruz, M. (2017). Adios, California. AltRight. Retrieved from https://web.archive.org/web/20180501012853/https://altright.com/

Salazar, P. J. (2018). The Alt-Right as a community of discourse. Javnost, 25(1?2), 135'143. https://doi.org/10.1080/13183222.2018.1423947

Sonnad, L., Squirrell, T. (2017, October 30). The alt-right is creating its own dialect. Here's the dictionary. Quartz. Retrieved from http://qz.com

Strickland, P. (2017, August 11) Alt-right rally: Charlottesville braces for violence. Al Jazeera. Retrieved from http://www.aljazeera.com

Strömbäck, J., & Esser, F. (2009). Shaping politics: Mediatization and media interventionism. In K. Lundby (Eds.), Mediatization: Concepts, changes, consequences, (pp.205:223). New York, NY: Peter Lang

Tarrow, S. G. (2011). Power in movement: Social movements and contentious politics, revised and updated third edition. Power in Movement: Social Movements and Contentious Politics, Revised and Updated Third Edition, 1?332. https://doi.org/10.1017/CBO9780511973529

What is 4chan? (n.d.) Retrieved from http://www.4chan.org

Wilson. J. (2015, 19 January). Cultural Marxism: A uniting theory for rightwingers who love to play the victim. The Guardian. Retrieved from https://www.theguardian.com/

Workman, K., Kannapell, A. (2015, June 18). The Charleston church shooting: What happened. The New York Times. Retrieved from http://www.thenewyorktimes.com

Appendix A

By the Eternal

Gregory Hood urges Donald Trump to embrace his inner Andrew Jackson.

'Badass' is a word that has been all but destroyed by the media, used to refer to fat feminists or limp wristed cucks who play tough guys on TV. In its true meaning, it refers to a man like Andrew Jackson, a leader comfortable with violence, jealous of his honor, and fiercely proud of his country. Jackson personifies something essential in the American character, especially the Scotch-Irish core that served as the American vanguard, and that, not coincidentally, is the core of the Trump coalition. The Scotch-Irish ethos is republicanism in its best sense, constituting a healthy distrust of decadent elites and a willingness to personally fight for one's faith and people.

A year ago, we were raging about how Andrew Jackson would be taken off the $20, as affirmative action was coming to currency. Now, the President of the United State is laying a wreath at Old Hickory's grave and openly appealing to the Jacksonian legacy.

Of course, he's not quite going far enough. As this is written, one judge after another is contemptuously tossing aside President Trump's travel bans on people from certain Muslim countries. The rationale for these judicial diktats is that Trump's executive orders are actually an attempt to impose a 'Muslim ban.' Thus, the statements of Trump and his advisors during the campaign are suddenly more relevant to these judges then the actual text of the executive order.

Obviously, the judges are driven by purely ideological considerations. As the SJW?s in black robes invent one newly invented 'right' after another with ever more convoluted Talmudic reasoning, it's hard not to have contempt for the whole process. Even listening to them appeal to the 'law' or to the 'Constitution' is obscene.

Forget the politics of it. Who can believe the Constitution shows that the Founding Fathers always intended for homosexual marriage to be imposed on every state by the order of the federal government?

Here, we see the rotten fruit of John Locke and John Rawls, after being dumbed down to serve an affirmative action culture. The state is conceived as a 'neutral' arbitrator that assigns bundles of rights to every upright biped on the planet, without regard or to culture, capability, or, at this point, even citizenship.

Never mind that John Locke and John Rawls themselves would not have agreed with it. (Even Rawls, by today's standards, was practically Pat Buchanan.) Defining the state as a legal framework supposedly 'neutral' to considerations of race and culture has now led to a regime that imposes a

plethora of special rights, restrictions and codes of behavior depending on what ethnic or sexual category each claimant fits into. Indeed, we are now at the point where even the distinction between citizen and non-citizen, already regarded as 'offensive' by the media, is not held to be important by the courts themselves. We are on the brink of the courts discovering an affirmative right to immigrate to the United States by every person on the planet. Somewhere in Valhalla, Carl Schmitt is chortling.

Again, forget the politics or your own views on each issue. The courts magically discovered 'rights' to abortion and gay marriage. They declared 'diversity' is a 'compelling state interest' requiring government discrimination against whites. They threw out Proposition 187, declaring that illegal immigrants had a 'right' to welfare from the very government whose laws they violated, and sealing the fate of the once great state of California. They destroyed American schools and mandated integration on reasoning that is now widely admitted to be faulty. Now, we are told hostile foreigners have a right to come into the country because Trump's campaign advisers said things that made them feel sad. The attempt by the courts to essentially deny the political and impose a 'neutral' regulatory structure has actually made the courts more political than ever. One dissent in the recent flurry of rulings, that actually (and reluctantly) admits President Trump does have the power to exclude aliens from this country, nonetheless engages in some virtue signaling against him. Perhaps the judge felt this was necessary so the dissent could even be written.

Judge Bybee writes, obviously referring to Trump: The personal attacks on the distinguished district judge and our colleagues were out of all bounds of civic and persuasive discourse-particularly when they came from the parties. It does no credit to the arguments of the parties to impugn the motives or the competence of the members of this court; ad hominem attacks are not a substitute for effective advocacy. Such personal attacks treat the court as though it were merely a political forum in which bargaining, compromise, and even intimidation are acceptable principles. The courts of law must be more than that, or we are not governed by law at all. But who can deny that is precisely what the court is? Of course a court is 'merely a political forum.' The courts are transparently political institutions. Each presidential elec-

tion, the nation as a whole admits this when each side screams 'this is the most important election ever' because 'the Supreme Court is at stake.'

The Left has turned justices like Ruth Bader Ginsburg into folk heroes for its side not because she is a great legal mind, or even a mediocre one, but because she is a transparent activist who will rubber stamp whatever ruling serves her agenda. Once there is a progressive majority on the court, we can take for granted it will use 'critical legal theory' that will characterize 'racist hate speech' (like citing accurate statistics about crime or group differences in intelligence) as no longer protected by the First Amendment. And as it stands now, there are no members of America's founding stock (WASP's) on the Court.

The American True Conservative insistence that 'the Constitution' has some magic power to freeze politics in place and prevent any infringement on what the Founding Fathers would have called our 'ancient rights as Englishmen' can no longer be accepted as a claim made in good faith. When we hear someone saying it, it's like hearing Evan McMullin ramble on that restricting immigration is 'white nationalism' or 'not our values.' It's a transparent attempt to exploit nostalgia to get Boomer-cucks and hopeless rubes to work against their own interest. It is impossible the Beltway Right hacks making these 'arguments' don't know exactly what they are doing.

The courts are not a check on power - they are a vehicle by that power can be exercised. The law is always a tool of power, indeed, the essential tool of power. The only thing that will stop power is an opposing power - and what we call 'liberty' is simply that space that exists within contending powers.

Power, like nature, abhors a vacuum. Liberty can only be guarded by concrete force. And the election of Donald Trump, rather than chastising the Left, is only encouraging the left wing managerial elite to seek even more unlimited control over property, speech, and social behavior.

Obviously, the whole sale abolition of immigration controls is a huge step in this effort. Salon recently opined that Open Borders is the best way to defeat Donald Trump. Salon is not wrong.

For that reason, it is time for President Trump to truly channel his inner Andrew Jackson. The courts have made their rulings - let them enforce it.

There are two ways to throw down the gauntlet. Originally, under the founding vision, the judiciary

branch was supposed to be the weakest branch of government. Now, it seems like it is the only one that matters. President Trump can take action against this by using the bully pulpit (now enabled by his promised biweekly rallies) to push Congress to strip courts of jurisdiction over immigration issues.

If we want to play legal games with the Constitution, consider Article III, Section 2, Clause 2 of the magic parchment: In all the other Cases before mentioned, the supreme Court shall have appellate Jurisdiction, both as to Law and Fact, with such Exceptions, and under such Regulations as the Congress shall make. The right to determine who can and cannot enter a country is fundamental to sovereignty. As the Supreme Court will soon have a (nominal) conservative majority, Trump has to push Congress to act now and restore control over this most critical of issues. This is the approach urged by some of the less destructive 'true conservatives' such as Daniel Horowitz at Conservative Review.

The second approach would be bolder and doesn't necessarily preclude doing the former tactic. As he did at his rally in Nashville, President Trump should simply read the laws that give him the obvious power to regulate immigration issues. He should order the executive branch to enforce all existing immigration laws as strictly as possible. He should end the DACA amnesty. And he should simply order the executive branch to stop processing visas from these countries. Force a constitutional crisis - let the judiciary try to stop him. Ultimately, such a crisis would go to the Supreme Court, where President Trump would likely win anyway. But the deeper purpose is to delegitimize the judiciary in the public eye, much as the 'Dishonest Media' or 'Lying Press' has been delegitimized.

And this isn't part of some dire scheme to hollow out American institutions - it's to expose the hollowing out that has already taken place. Such a battle is necessary and essentially if leftist activists masquerading as judges aren't just going to throw out everything President Trump does anyway. Leftists are already screaming for President Trump's impeachment. Were Democrats in the majority, it is unquestionable such an action would have taken place. Impeachment is not about checking illegality or wrongdoing - it's a political act, much like a 'vote of no confidence' in a Prime Minister usually means that a PM has lost political support, no committed some crime. Bill Clinton's overcoming his impeach-

ment in the 1990's was a spectacular illustration of this principle. And it is time to view the courts the same way - as merely political institutions, battlegrounds for competing actors seeking power.

As it stands now, the American Left already regards the law and the courts this way. They never accept a verdict until it is the one they want. The American Right, meanwhile, throws up its hands and meekly goes along with defat, professing loyalty to some abstract process rather than its actual constituency. They fight for procedure, rather than fighting for the essence. It's not a question of hoping President Trump does what we want. It's a question of what President Trump needs to do in order not to be rendered impotent within the first months of his term. If President Trump wants to accomplish anything, he needs to look to Andrew Jackson rather than to Buckley, and he needs to welcome this fight. What Andrew Jackson said of the Bank of the United States must now be applied to the courts: You are a den of vipers and thieves. I have determined to rout you out, and by the Eternal, I will rout you out!

Appendix B

Adiós, California

A Californian boomer offers her first-hand account of the state's infamous demographic shift and the frustrations over her generation's inability to stop illegal immigration. Submitted by Melinda Santa Cruz

As far as I was concerned, that was a word that described my parents' generation. To my surprise and thanks to the Alt-Right, I've learned that I do fit the boomer demographic. I was born in 1958 and was too young to have experienced the hippie movement. My parents were Republicans and wouldn't have allowed such behavior anyway.

While the generation in between was smoking dope, getting high on LSD, living in vans, creating social unrest, and protesting the Vietnam War, my family was tucked safely away in an upper-middle class neighborhood in Orange County, California. Throughout the war, I wore a metal bracelet with the name of a POW, Major Don Lyon. No one told me to wear it. I never received a lecture on how to be a patriot. I simply knew what my family stood for. The fact that my grandfather had been an officer on the Santa Ana police force or that my fa-

ther had served as a U.S. Marine during the Korean War probably had something to do with it. But without any brothers, I spent most of my childhood surrounded by ballerina figurines and collector's dolls. I had little knowledge about anything manly such as risking your life for your fellow man or service to your country.

Sometime around 1970, the demographics in California started to change. By then we had moved to a five-acre ranch in North San Diego County and the populace around us was growing darker in terms of skin color. There weren't any blacks around. They stayed in their own communities in Los Angeles. The people who were infiltrating our society were Mexican Indians because the upper class in Northern Mexico discriminated against them.

Ours was the last house before you went down to the Rincon Indian Reservation. There was no police or fire service; in the tradition of the Scots-Irish, you simply took care of business on your own. I was about fifteen and home alone when a group of illegal aliens got frighteningly close to the house. I took out my dad's twelve-gauge shotgun, opened the window, and placed it on the sill. The group immediately took off.

My mother did hire an illegal couple to work on the ranch. Esperanza cleaned the house and took care of my sister and me while my mother worked. Jesus fed the horses and did whatever else needed to be done on the property. We knew they were illegal and it royally pissed me off because they were taking advantage of our good will without having regard for our laws. But they weren't bad people overall. They were peasant farmers from Guadalajara without any opportunity to thrive in their own country. I did have respect for the fact that, unlike illegal aliens today, they would have sacrificed their lungs to become Americans and were prepared to throw their allegiance to Mexico under a bus.

During those years they learned English while I learned very little Spanish. Jesus insisted that I speak to him in English even if he didn't understand because he wanted to assimilate. He and I would sit under an oak tree with his English-Spanish dictionary and I would teach him how to pronounce the words while the warm smell of Esperanza's hand-made tortillas wafted around us.

One day, I came home from school to find my mom and Esperanza in a state of chaos. Jesus had been driving along Valley Center Road when a po-

lice officer had pulled him over and discovered a warrant. An old girlfriend claimed that they'd had a baby together and that he owed child support. Since he was an illegal alien, he was deported on the spot. I cried, thinking we would never see him again, I had really appreciated his thoughtfulness. On my birthday, he'd given me a five-dollar bill as a gift when he had none to spare. Imagine my surprise when I woke up three days after he'd been picked up to find him tossing a flake of hay into a horse's stall. His old girlfriend had confessed that the child was not his in Mexican court. He had then skipped back across the border without a hitch. It was then that I realized just how much open borders could screw us up. 'Come on in. Food Stamps for everybody . . .' By the late 1970s, my family had moved back into town. The plan was to stay in a cramped house in a working-class neighborhood until the kids graduated from high school. Then the family would move to a better neighborhood. In the 1990s, my mother saw an old friend from our days on the ranch.

'Do you still hire the wetbacks?' she asked.

'Oh hell no,' the friend said. 'We are terrified. They come across the border with prison tattoos and guns. We wouldn't dare.'

The days of having peasant farmers crossing the border for a better life were over. We now had hardcore criminals committing murder and rape.

My grandparents still lived in Santa Ana in a house built by my great grandfather in the 1930s. The neighborhood seemed to change overnight as senior citizens were replaced by Mexicans and even Central Americans from places like El Salvador. My grandparents' property was robbed twice. The first time, the thieves got the heirloom jewelry that I was to inherit, and a shoebox full of silver dollars. The second time, a case full of dusty cassette tapes was stolen from the garage. An elderly woman that my grandparents has known for decades woke up in the middle of the night to find a Mexican man on top of her. She later said that it was only by the 'grace of God' that she was able to hold him off.

I started to write letters to the editor about the illegal alien problem in 1984. The letters were published in the North County Times and the San Diego Union Tribune but didn't get much attention.

My grandfather was the only white person left in the town where he had lived for sixty years. He was known as 'Mr. Orange County' because of his participation in civic affairs. When mariachi mu-

sic shook the walls of his small house one day, he waved his old police pistol around and declared, 'I'm going down and I'm taking the Mexicans with me.' He was eighty-eight and my uncle came from Los Angeles to take the gun away.

At the time, I wrote my grandfather off as a crazy old man, but now I see that he was traumatized by what was happening around him. The same thing is happening to me in North San Diego County today.

My daughter was preyed upon by Mexican gang-bangers who thought she was one of them. Her father was Spanish, not Mexican. She was called a 'disgrace to her race' and told to stop 'acting white.'

'What are you talking about?' she asked. 'I am white.'

Her European ancestry made no difference to them. I transferred her to another high school only to find that Mexican students were being bussed into that white neighborhood. I had to drive her there every morning because there was no bus for her, but there were two for them. And the globalists wonder why we voted for Trump? This would be one reason. Mexican women have intentionally hit the back of my legs with baby strollers and nudged me out of the way while I'm shopping at Ross. I've made up my mind that if one of them does it again, I'll make my opinion known and accept the consequences. Because of acts like this, boomers couldn't race to the polls fast enough to vote for Ronald Reagan. He was to be the savior that would

make America white again. We were all devastated when he proved himself a traitor. Against our wishes, he granted amnesty to six million illegal aliens and started the current immigration trend of 'We'll allow the ones already here to stay and secure the border for the future.' Boomers voted for Pete Wilson for governor and for Proposition 187 to rectify the situation. Wilson did nothing and a judge overturned the voter-approved measure. Apparently, it is unconstitutional for taxpayers to not want to pay for the welfare benefits and education of freeloaders. Trump was our great white hope and even he has let us down in favor of the Hispanics. Who cares if Pepe the Pizza Maker gets deported? So what if an MS13 gang member kills someone in his own neighborhood? Most of us will never be the victim of that level of violence. The real threat is DACA (Deferred Action for Childhood Arrivals) and what they are doing to the job market for professionals. So far, thirty-nine Dreamers have gone to work for Facebook alone, and Mark Zuckerberg is fighting for more. If this keeps up, where will the truly qualified white professionals going to work? What has happened to our merit-based American dream? The Alt-Right tends to blame the boomers for the whole immigration mess. Trust me when I say that many of us did everything we could to stop it. Every politician who promised to fix the problem got elected into office and then proved himself to be a traitor. We were against an unstoppable tide from the beginning. Now, the only thing left is to say goodbye.

Social Sciences

Peer Culture as Moroccan-Dutch Youths' Space of Resistance: "We're Kinda Creating Our Own Little Thing As We Weren't Accepted in the First Place"

Domiziana Turcatti

Abstract

This critical ethnography investigated the role of peers in the lives of 30 Moroccan-Dutch youths aged 17-26 living in Amsterdam and Rotterdam. The research aimed to identify (1) the processes underlying peer group formation; (2) the role of peer culture in Moroccan-Dutch youths' identity-formation; (3) the kinds of support that peers provide Moroccan-Dutch youths. Framed by a post-colonial theoretical framework merging peer culture, peer network, social support, and acculturation theory, the study is based on 110 hours of fieldwork conducted between January 2017 and May 2018. Challenging dominant constructions of Moroccan-Dutch youths and their peers, this research reveals that the formation of diverse peer groups not always including native Dutch is underlined by segregation, an exclusive construction of Dutch identity, and a search for understanding. Furthermore, peer groups become sites of support and spaces where the youths forge their identities, while creating a peer culture valuing success, community, and diversity. However, peers are not necessarily a source of opportunities, and discrimination and inequality threaten such a peer culture. While demonstrating the need to decolonize research and theory through the youths' critical thinking, the study indicates the urgency for evidence-based policies dismantling structural inequality, encouraging supportive networks and intercultural respect

Keywords and phrases: Moroccan-Dutch youth, peer culture, identity formation, acculturation, peer support, peer social capital

1 Introduction

This research investigates the meaning-making that Moroccan-Dutch[1] youths and their peers collectively construct and deploy to deal with everyday challenges. The youths face multiple barriers to entering the labour market and to achieving high-status occupations (Turcatti & Scott, 2015; Turcatti, 2017d; Turcatti, Berger & Kruimer, 2017). These obstacles are exacerbated by an educational system disfavouring the youths' academic achievement (Turcatti, 2017b; Turcatti, 2018). Furthermore, the youths need to confront hegemonic constructions of their heritage culture and religion as the 'Other', which make the acculturation process - namely the negotiation of heritage and Dutch culture - more difficult (Turcatti, 2017a; Turcatti, 2017e; Turcatti, 2017f). While such structural forces have been documented, little research has focused on the youths' strategies in navigating Dutch society. Hence, this research focused specifically on the role that peers play in Moroccan-Dutch youths' navigation in Dutch society.

Scholars in the Netherlands have overwhelmingly researched the negative influence of peers on Moroccan-Dutch youths' well-being with a quantitative approach (Demanet & Van Houtte, 2012; Dekovic, Wissink & Marie, 2004; de Jong, 2012; Weerman, Bernasco, Bruinsma & Pauwels, 2015; Wissink, Dekovic & Meijer, 2009). However, the role of peers in providing support and as channels of social capital for Moroccan-Dutch youths seems to have been overlooked. Furthermore, it seems that the role of peers and peer culture as influencing the construction of Moroccan-Dutch youths' personal identity has been a marginal topic in the literature. Quantitative research has focused more on how contact with co-ethnic peers and rejection by native Dutch peers negatively influence the acculturation process of Moroccan-Dutch youths (De Vroome, Verkuyten, & Martinovic, 2014; Azghari, van de Vijver & Hooghiemstra, 2017).

The fact that research has predominantly focused on the negative influences of the youths' peer networks is not surprising, given that scholarship on Moroccan-Dutch youths has mostly focused on their failures, and only recently have scholars started to investigate their successes (Turcatti, 2017d). However, such focus is problematic insofar as it contributes to the pathologization and

---

[1]The term Moroccan-Dutch youths is used to identify the youths born in the Netherlands and who have at least one parent who was born in Morocco. Hence, the term Moroccan-Dutch youths is used to identify, following Rumbaut's (2004) classification, second-generation Moroccans (the youths born in the Netherlands and having both parents from Morocco) and 2.5 generation Moroccans (the youths born in the Netherlands who have one Moroccan parent and one Dutch parent). However, these two terms will not be used, unless to specify their position in Dutch society as children of migrants. This is because I believe they are politically charged insofar the terms 'second or 2.5 generation Moroccans' label the youths as migrants and do not recognize their Dutch belonging. Thus, the term Moroccan-Dutch will be used as it seems a more inclusive one that acknowledges and values both the heritage and Dutch culture.

blame on Moroccan-Dutch youths (and their peer networks) for their lack of integration according to Dutch media and policy discourse (Vasta, 2007). This inevitably leaves unchallenged the structures of inequality constraining the opportunities available to them. More specifically, a focus on the negative influence of peers in the lives of Moroccan-Dutch youths prevents from discovering the resources that peers might provide and how that could be strengthened through intervention. Furthermore, it leaves unknown how peers and their peer cultures might become important spaces for defining how they fit in a society characterized by ethnic hierarchies.

In light of such academic and socio-political context, I conducted a critical ethnographic research[2]. The research was interpretive insofar it aimed to uncover how the participants make sense of their peer relationships from their own perspective. The research was also critical, in the sense that it examined the meaning-making of the participants and their peers 'through the lens of power, prestige, privilege, and authority' (Harrowing et al., 2010, p. 242). This methodological approach was adopted because scholars have previously tended to conduct quantitative, longitudinal studies about the youths' position in Dutch society. Hence, this approach was necessary to address the often-dismissed narratives of Moroccan-Dutch youths and to 'contest the cumulative impact of sustained inequalities that produce marginalization and privilege' (Weis & Fine, 2012, p. 175). The research questions that guided the research are:

1. Who are the peers of a group of Moroccan-Dutch youths living in Amsterdam and Rotterdam?

2. In what ways do peer relationships become an expression of 'culture' or a space for Moroccan-Dutch youths' identity-making?

3. In what ways do peers provide supportive mechanisms to the group of Moroccan-Dutch

youths living in Amsterdam and Rotterdam participating in the research?

The research was conducted under an interdisciplinary, postcolonial theoretical framework merging theory from sociology, anthropology and psychology. The research is indeed framed under peer culture theory (Corsaro & Eder, 1990), peer network theory (Stanton-Salazar, 2004; Gibson, Gandara & Koyama, 2004), social support theory (House, 1981), stress theory (Lazarus & Folkman, 1984; Masten, 1997), and identity and acculturation theory (Berry, 2005; Schwartz et al., 2006). Methodologically, the research is based on 110 hours of fieldwork conducted between January 2017 and May 2018 with 30 Moroccan-Dutch youths (18 males, 12 females) aged 17-26 living in Amsterdam and Rotterdam. During the fieldwork, qualitative data was collected in the form of one focus group, paired and individual semi-structured interviews, life histories, informal conversations, participant observations, life history maps, social support maps, and peer maps.

This research aimed to contribute to the existing body of knowledge and to the theoretical understanding of the role that peers might play in the lives of urban youths from migrant families. Furthermore, the research aimed to advocate in favour of a shift away from a focus on deviance and disenfranchisement to one that investigates the strengths of the youths. Indeed, the research is grounded in the belief that such a shift in focus allows to locate 'failure' and 'deviancy' in broader structures of oppression which, through the appropriate policies and practices, can be addressed, softened, and potentially eliminated. This approach was ultimately chosen in order to identify peer structures and networks that can be strengthened and to spot structures limiting the functioning of Moroccan-Dutch youths as citizens and as human beings that can be changed.

---

[2]Critical ethnography is the term generally employed to indicate a specific way of collecting and analysing qualitative data (Carspecken, 1996; Madison, 2012; Tracy, 2013). Critical ethnography examines 'culture through the lens of power, prestige, privilege, and authority' (Harrowing, Mill & Spiers, 2010, p. 242). Researchers conducting critical ethnography tend to share the following set of value orientations: research should serve as cultural and social criticism; the inequality of some societal groups is linked to the privilege enjoyed by other societal groups; oppression is reproduced and has many faces; researchers can reproduce inequalities and need to employ active reflexivity in the attempt to not do so; critical researchers have the ethical responsibility to address structures of oppression, advocate against them and employ the research in favour of social justice and change (Fine, 1994; Carspecken, 1996; Madison, 2012; Tracy, 2013).

2 Historical and Socio-Political Context of the Netherlands

By drawing from academic and government literature, this chapter presents the context in which the research is situated. The first section focuses on the Moroccan migration to the Netherlands. The second section describes the position of Moroccan-Dutch youths in Dutch society.

## 2.1 From Morocco to the Netherlands: Migration & Settlement in A Diverse Society

Approximately 22.1 percent of the Dutch population has a migrant background (CBS, 2016)[3]. Moroccan-Dutch youth belongs to what the government defines the second-largest 'non-Western' 'ethnic minority' group (CBS, 2016)[4]. Within this group, 56 percent are individuals who are born in the Netherlands from at least one Moroccan parent (CBS, 2016).

Moroccans started to migrate to the Nether-

lands in the sixties (Figure 1). At that time, the Netherlands faced a shortage of labour in the industrial sector, which motivated the Dutch government to seek low-skilled workers in countries such as Morocco and Turkey (Tesser & Dronkers, 2007). Thus, Moroccans who moved to the Netherlands as guest workers mostly had low levels of education and came to occupy a lower socio-economic position in comparison to their Dutch counterparts (Crul & Doomernik, 2003; Tesser & Dronkers, 2007).

The position of Moroccan guest-workers became more precarious following the reconstruction of the Dutch economy in the 1980's, since the demanded level of qualification became much higher (Tesser & Dronkers, 2007). This led to a sharp increase in the rate of unemployment among Moroccan guest-workers (Tesser & Dronkers, 2007). They settled down in urban areas where the most jobs and social housing were available (Peters, 2011; van Praag & Schoorl, 2008) and started to reunify their family in their new home country (Tesser & Dronkers, 2007).

Figure 2.1   *Immigration of persons with Turkish or Moroccan citizenship, 1969-2006*
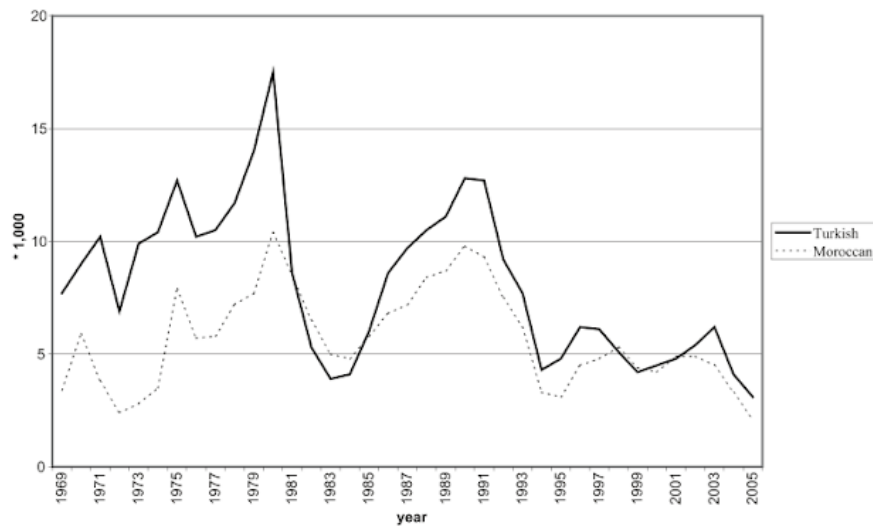
Source: Statistics Netherlands

---

[3]The government classifies the inhabitants with a migrant background as those who have a 'Western' background (9.8 percent of the whole population) and those who have a 'non-Western' background (CBS 2016). The major groups that the government identifies in the latter category are Turkish, Moroccan, Surinamese, and Antillean.

[4]The term 'ethnic minority' is used in governmental reports to indicate individuals who are migrants or have a migrant background and, as such, cannot be 'natives' or 'native Dutch'. I will try to limit the use of the terms 'ethnic minority' and 'natives' or 'native Dutch', unless these are necessary to highlight the differences in positionality between native Dutch youths and Moroccan-Dutch youths. This is because these terms are politically charged in the sense that they establish asymmetrical power relations between those who are the majority and those who are not, and between those who are considered natives and entitled to the belonging and those who are not (Wekker, 2016)

Migration of Persons with Turkish or Moroccan Citizenship to the Netherlands, 1969-2006 (Crul & Herring, 2008)

Moroccan migrants' settlement ' in conjunction with the settlement of migrants from other countries[5] ' led the government to initiate a new integration policy in the eighties (Pennix, 1979; Essed, 1996). However, this was also the time when Moroccan immigrants started to be perceived as a 'problem', insofar as their settlement meant that they needed proper housing, schooling, and mosques (Essed, 1996 p. 17). Yet Essed (1996) emphasizes that Moroccan migrants were more than a 'problem'. They became a 'threat' precisely because they not only needed facilities, but because 'they openly claimed equal housing, equal schooling, and equal [emphasis added] job opportunities' (Essed, 1996, p. 17).

2.2 The Moroccan-Dutch Youth in Dutch Society

2.21 The Socio-Economic Position of Moroccan-Dutch Youths

In the past two decades, academic and government literature reports that Moroccan-Dutch youths have made improvements in the educational and labour domains (Crul and Doomernik, 2003; Tesser & Dronkers, 2007; Forum; 2010; CBS, 2012; CBS, 2016). In the educational realm, the scores of Moroccan-Dutch youths in the Cito exam[6] and the percentage of Moroccan-Dutch youths attending university have increased (Crul and Doomernik, 2003; Tesser & Dronkers, 2007; Forum; 2010; CBS, 2012; CBS, 2016). Similarly, Moroccan-Dutch youths' performance in the labour market has improved as they come to occupy

higher-status positions (Crul and Doomernik, 2003; Tesser & Dronkers, 2007; Forum; 2010; CBS, 2012; CBS, 2016).

Nevertheless, the socio-economic position of Moroccan-Dutch youths is characterized by a polarization between those who are improving their socio-economic position and those who lag behind (Crul & Doomernik, 2003; Slootman, 2014). For this reason, on average they do not perform as well as the so-called native Dutch counterpart and other so-called 'minority' groups in education and in the labour market (Tesser & Dronkers, 2007; Crul and Doomernik, 2003; CBS, 2016; CBS; 2012; Forum; 2010). On average, governmental and academic sources report how Moroccan-Dutch youths are under-represented in higher tracks of the high school educational system, which instead tend to be overrepresented by native Dutch youths, and Moroccan-Dutch youths tend to leave school at an earlier age more often (Tesser & Dronkers, 2007; Crul and Doomernik, 2003; CBS, 2016; CBS; 2012; Forum; 2010).

Furthermore, in the past two decades, Moroccan-Dutch youths have tended to occupy lower-status positions (Gracia, Vázquez-Quesada & Van de Werfhorst, 2016) and, as Figure 2 shows, they have been suffering from extremely high rates of unemployment when compared to other societal groups (Tesser & Dronkers, 2007; Crul and Doomernik, 2003; CBS, 2016; CBS; 2012; Forum; 2008; Forum; 2010). From Table 1, it can be noticed that in 2015 the average unemployment rate of the Moroccan-Dutch youth amounted to twenty-two percent whereas the unemployment rate among native Dutch amounted to six percent (CBS, 2016).

---

[5]Moroccans were not the only one who started to migrate to the Netherlands after the Second World War. Indeed, the first immigrants after the Second World War were from Indonesia ' a former Dutch colony ' and were followed by Moluccans (Essed, 1996). Starting from the sixties, families from Italy, Greece, Turkey and Morocco began to migrate to the Netherlands (Essed, 1996). In the seventies and eighties, people from the former colony of Suriname, from the Dutch Antilles (Essed, 1996) and from Latin America (Sandoval, 2008) also started to move to the Netherlands. Later in the nineties, the Netherlands saw an increase of asylum seekers and of Ghanaians (Jloosterman, Ruisinovic & Yeboah, 2016), which was followed by the arrival of migrant workers from central and eastern Europe (Engbersen, Leerkes, Grabowska-Lusinska, Snel & Burgers, 2013).

[6]The Cito exam is the test that children take in their eighth year of elementary school which examines students' math and language ability. The results of the Cito exam together with the school advice determine the type and level of high school they will attend (Stevens, Clycq, Timmerman & Van Houtte, 2011). A more detailed explanation of the Dutch educational system is provided in section 2.2.2
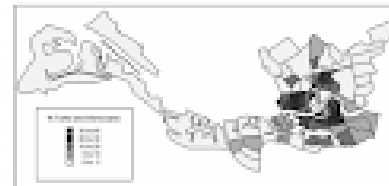
2.3.1 Unemployment (first quarter)
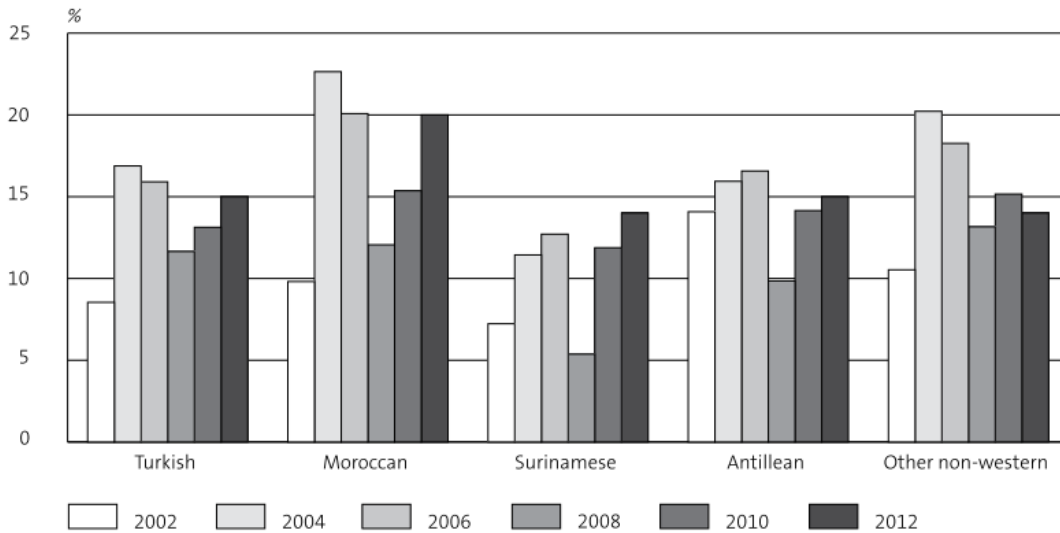
Source: Statistics Netherlands.

Variation of Unemployment Rates in the Years 2002-2012 among Youth Born in the Netherlands but of a 'non-Western' Background (CBS, 2012)

Regarding the housing situation, approximately two-thirds of Moroccan-Dutch youths live in the largest Dutch cities (Huijnk & Andriessen, 2016). In particular, Amsterdam and Rotterdam are the cities with the largest share of individuals of Moroccan background (Crul & Heering 2008). Figure 3 and 4 respectively provide a proxy of the areas where the descendants of Moroccan migrants live which tend to be highly diverse, inner-city neighbourhoods (Crul & Heering, 2008). Indeed, two-thirds of the Moroccans settled in areas characterized by at least sixty percent of 'non-Western' immigrants (Musterd & Fullaondo, 2008). These areas have been labelled as segregated because of the low percentage of native Dutch families (Muster & Fullaondo, 2008).

Population of Turkish and Moroccan Descent as a Percentage of the Total Population in Postal Code Areas of Rotterdam (Van Praag & Schoorl, 2007)

2.22 Moroccan-Dutch Youths' Challenges and Barriers to Social Mobility

Qualitative and quantitative research has shown that the Moroccan-Dutch youth, as a group, faces multiple, cumulative obstacles which limit the opportunities available to them in multiple domains of Dutch society (Crul & Doomernik, 2003; Crul & Heering, 2008; Heath, Rothon & Kilpi, 2008; Turcatti, 2017b). Following Heath et al. (2008), these obstacles might be generally categorized into four

---

[7]This has been explained by the fact that their families tend to lack the economic, social, and cultural resources to transmit to their children. For instance, they might not have the economic means to invest in extra-scholastic, (non)educational activities, they might not master the Dutch language, and lack knowledge about Dutch institutions and their functioning (Crul & Doomernik, 2003; Heat et al., 2008; Turcatti, 2017b). Furthermore, they might lack social capital (Crul & Doomernik, 2003) ' the social networks and relationships that allow access to resources that might benefit one's socio-economic status (Bourdieu, 1986) ' as well as the cultural capital (Gracia et al., 2016) ' the sociolinguistic standards and cultural behaviours (Bourdieu, 1986) ' that might benefit one's standing in society (Bourdieu, 1977).
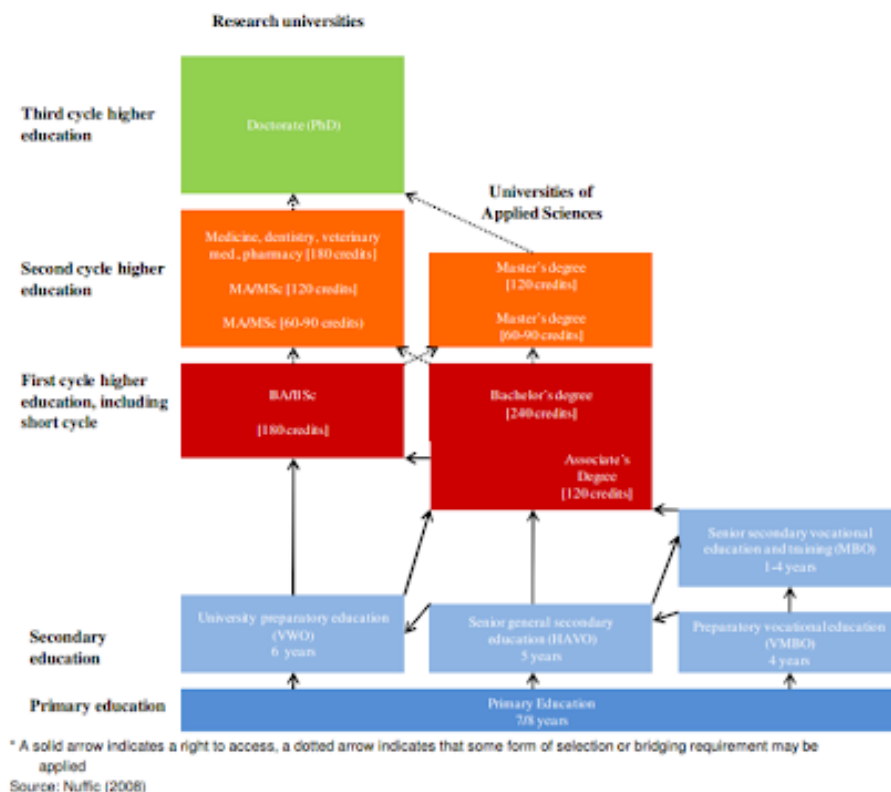
categories: (a) family-related obstacles, (b) ethnic segregation, (c) structural barriers, and (d) discrimination.

Coming from a lower socio-economic background is identified as one of the central challenges that the youths need to confront, since it reduces the chances of acquiring a high-status occupation and of accessing university[7] (Crul & Doomernik, 2003; Tesser & Dronkers, 2007; Gracia et al., 2016). These family-related obstacles are exacerbated by the challenges that growing up in a highly diverse, low income, segregated neighbourhoods pose (Crul & Doomernik, 2003; Crul & Heering, 2008; Heat et al., 2008). These neighbourhoods might lack many resources and opportunities, such as quality education and knowledge circulation (Heath et al., 2008), which are necessary for socio-economic mobility (Heath et al., 2008).

Furthermore, Moroccan-Dutch youths also need to deal with the structural barriers posed by the highly regulated Dutch educational system and labour market (van Zenderen, 2010). The Dutch educational system features an early tracking system that assigns pupils either to VMBO (prevocational or junior general secondary education), HAVO (senior general education) or VWO (pre-university education). As it can be seen from Figure 5, of these, only VWO gives direct access to university. Similarly, the Dutch labour market is highly regulated as for each occupation well-defined educational titles are required (Di Stasio & Van De Werfhorst, 2016; Van De Werfhorst, 2011). This means that there is a lack of 'high social mobility if one lacks qualification and did not undergo the 'right' institutional and educational path' (Di Stasio & Van De Werfhorst, 2016). This reduces the chances of Moroccan-Dutch youths who have left school earlier or have not pursued the 'right' educational path to access high-status occupations (Di Stasio & Van De Werfhorst, 2016).

Structure of the Higher Education System in the Netherlands (City of Rotterdam Regional Steering Committee, 2009)

Lastly, Moroccan-Dutch youths need to deal with institutional and everyday racism in many domains of their lives (Essed & Hoving, 2014; Wekker, 2016). Research has largely acknowledged the presence of discrimination in the labour market and the detrimental consequences it has on the upward socioeconomic mobility of Moroccan-Dutch youths (Andriessen, Nievers, Dagevos & Faulk, 2012; Gracia

et al., 2016). In terms of education, however, academic literature seems to deny the role that discrimination plays in the Dutch educational system (Van De Werfhorst & Van Tubergen, 2007). However, qualitative studies have revealed the presence of teachers' discriminatory practices that negatively influence Moroccan-Dutch youths' schooling experience and their academic achievement (Stevens et al., 2011; Turcatti, 2018).

2.2.3 Depictions of the Moroccan-Dutch Youth in Dutch society

In Dutch society, it seems that Moroccan-Dutch youths also need to deal with cultural oppression and hegemonic representations that construct the youths as the pathologized, racialized Other (Essed & Hoving, 2014; Wekker, 2016). In Dutch media and political discourses, Moroccan-Dutch youths are often constructed as an object of problematization (Schinkel, 2013; Wekker, 2016) insofar they are depicted as socially and economically unintegrated and reluctant to do so, effectively dismissing if not denying the obstacles and lack of opportunities the youths confront (Vasta. 2007; Schinkel, 2013; Essed & Hoving, 2014; Wekker, 2016).

The tendency to negate the role that inequality plays in the lives of Moroccan-Dutch youths also characterizes policy and governmental discourse (Vasta. 2007; Schinkel, 2013; Essed & Hoving, 2014; Wekker, 2016). Since the second half of the nineties, governmental and policy discourse began to promote the idea that 'non-Western' 'ethnic minorities', and in particular Moroccan and Turkish ethnic 'minorities', did not achieve socio-economic emancipation, despite the efforts of the government (Vasta. 2007; Schinkel, 2013; Essed & Hoving, 2014; Wekker, 2016). The reasons provided were attributed to 'first, a misplaced tolerance for cultural difference on the part of the Dutch, and, second, some immigrants' deliberate refusal to embrace Dutch culture, language and values' (Vasta, 2007, p. 715). This discourse contributed to legitimize the shift from a policy encouraging multiculturalism to an assimilationist one discouraging the maintenance of religious and cultural heritage practices and not implementing measures to reduce inequalities (Vasta. 2007; Schinkel, 2013; Essed & Hoving, 2014; Wekker, 2016).

The way Moroccan-Dutch youths are depicted and perceived in Dutch society worsened since 9/11 and the murder of Theo Van Gogh[8]. These events, in conjunction with the spreading of Islamophobia and nationalistic sentiments throughout Europe (Buijs, 2009; Zaal, 2014), generated 'a preoccupation with the un-assimilability of the different and backward cultures of migrants' (Wekker, 2016, p. 55). These events also gave supposed legitimacy to claims about the inferiority of Moroccans' heritage culture and of Islam, as they are believed to promote values fundamentally incompatible and antithetical with those of tolerance and gender equality (Essed & Hoving, 2014; Goldberg, 2014; Wekker, 2016). These values are constructed as central to Dutch identity (Essed & Hoving, 2014; Goldberg, 2014; Wekker, 2016), which, according to Ghorashi (2016), is commonly constructed as 'based on colour, 'roots', and certain codes of behaviour that excludes difference' (p. 225). In this context, we might wonder what 'opportunities for selfhood' (Sirin & Fine, 2007, p. 151) are left for Moroccan-Dutch youths.

3 Theoretical Framework

This chapter presents the theories and literature used to frame the study. By drawing from Dutch and US scholarship, this chapter provides an overview of the literature (1) defining peers, peer relationships and peer culture; (2) addressing the role of peers in the construction of migrant youths' identities; and (3) theorizing the role of peers in the lives of urban youths from migrant families as providers of support.

3.1 Defining Peer Relations and Peer Culture Theory

3.1.1 Defining Peers, Peer Relations and Peer Groups

The term peers is used 'to refer to persons who occupy equivalent positions in an organization or social network' (Bank, 1997, p. 880). For youths, peers usually 'imply age mates, friends, acquaintances, and other individuals in close proximity and of equal status' (Gibson et al., 2004, p. 4). The relationships they establish among each other might

---

[8]Theo Van Gogh was a Dutch film-maker, who was killed by a young Moroccan-Dutch man after the publication of the short film Submission (Van Gogh, 2004), which attacked certain Islamic practices (Buijs, 2009).

be experienced as 'comfortable and rewarding' or 'alienating and hurtful', and may vary in strength (Gibson et al., 2004, p. 4).. As Granovetter (1973) theorized, the strength of the ties established with other individuals is a 'combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize it' (p. 1361).A peer group is used to indicate 'groups of two or more peers who are linked together by more [emphasis added] than their common identity level' (Bank, 1997, p. 880). The links between members of a peer group include 'contact, interaction, and positive, sociometric choices' (Bank, 1997, p. 880), where sociometric choices refer to mutual validation (e.g. you are my friend).

In the Netherlands, research on Moroccan-Dutch youths' peers and who they befriend is limited. Crul, Schneider and Leile (2013) conducted a quantitative research on Turkish-Dutch youths' friendships in Amsterdam and found that almost sixty percent of them engage in what they call 'ethnically mixed groups of friends' (p. 77) while only 30 percent of the native Dutch youth do so. Crul et al. (2013) defined such youths as the 'hinge generation,' namely the generation who is 'accustomed to moving between their own community and a variety of other ethnic groups from an early age' (p. 78). Following Crul's et al. (2013) study, this research tried to identify who the peers of a group of Moroccan-Dutch youths are in terms of background, but also in terms of other social identities and social locations. The research sought to uncover how their peers are classified and described, the attributes that the youths value in their peers, and what kinds of relationships they establish with them.

3.1.2 Peer Culture Theory

Peer culture is the term originally coined by William Corsaro (1988) to indicate 'a stable set of activities or routines, artefacts, values, and concerns that children [and youths] produce and share in interaction with peers'[9] (Corsaro & Eder, 1990,

p. 197). Hence, peer culture is 'the descriptive and evaluative meanings that peer groups assign to behaviours and relationships' (Bank, 1997, p. 880), with the meanings never static but always constructed, challenged or consolidated through interaction (Corsaro, 1988; Corsaro & Eder, 1990). According to Corsaro (1988), peer groups provide an important context of socialization that is distinct from the familiar one and which might give life to an autonomous peer culture (Corsaro, 1988; Corsaro & Eder, 1990).

As limited research has explored whether the interactions among Moroccan-Dutch youths engender a peer culture, this research investigated whether a group of Moroccan-Dutch youths address their peer relationships in terms of what might be labelled as attributes of culture (e.g. norms, values and belief) and the ways these are connected to the external context and power dynamics.

3.2 Peers and Peer Culture: Spaces for Identity Construction

Before presenting the existing literature and theory on peer relationships and the construction of urban youth identity, it is necessary to explain that the terms identity and culture are approached from an anti-essentialist standpoint. These terms were adopted by post-colonial theorists such as Homi Bhabha (1994) and Stuart Hall (1991), who consider identity and culture to be unfixed and essential entities as existing 'out there', territorially confined and folk-bound. On the basis of this understanding of identity and culture, this section will present the theory and literature (1) on the relationship between peers and personal identity formation, and (2) on the relationships between peers and cultural and social identity formation.

3.2.1 Peers, Peer Culture and Personal Identity Formation

Peer groups and their peer culture become increasingly important for adolescents and youths

---

[9]From this definition of peer culture, it can be seen that Corsaro (1988) espouses Geertz's (1973) interpretive and ideational theory of culture (Keesing, 1974). According to Geertz (1973), 'though ideational, [culture] does not exist in someone's head; though unphysical, it is not an occult entity' (Geertz, 1973, p. 10). Indeed, culture becomes both a shared 'frame of mind' and a 'patterned conduct' (Geertz, 1973, p. 10). According to Geertz (1973), 'behaviour must be attended to, and with some exactness, because it is through the flow of behaviour ' or more precisely, social action - that cultural forms find articulation' (Geertz, 1973, p. 17). More specifically, by using the term behaviour, Geertz (1973) refers to what he defines as 'control mechanisms'. Hence, this research ' as Corsaro (1988) did ' also endorsed Geertz's interpretive theory of culture to discover the meanings that Moroccan-Dutch youths share with their peers and how such meanings regulate, govern and are embedded in their everyday activities.

who need to reduce their dependency in psychological terms from adults, allowing them 'to establish a unique and autonomous identity different from that of one's parents' (Haayen, 2016, p. 74). Hence, peer groups and their culture influence the construction of youths' personal identity (Haayen, 2016), which refers to 'the goals, values, and beliefs that an individual adopts and holds' (Schwartz et al., 2006, p. 6). This might be particularly important for youths belonging to a stigmatised group who might internalize diminishing hegemonic characterizations and therefore underperform. This process ' named 'stereotype threat' by Claude Steele and Joshua Aronson (1995) ' was found to have detrimental consequences, insofar it can lead to underperformance in different domains, ranging from education, labour or everyday activities such as playing video games (Schmader, 2010; Kinias & Sim, 2011; Spencer, Logel & Davies, 2016; Vermeulen, Castellar, Janssen, Calvi & Looy, 2016).

3.2.2 Peers, Peer Culture and Cultural and Social Identity Formation

Peers and peer culture also influence and offer a space for the construction of the youths' social identities (Vigil, 2004; Shin et al., 2007; Slootman, 2014; Haayen, 2016;), which 'refe[r] both to (a) the group with which one identifies, including its self-identified ideals, mores, levels and conversions and (b) the extent to which this identification leads one to favour the 'ingroup' (i.e. the group to which one perceives oneself as belonging) and to distance oneself from 'outgroups' (i.e. groups other than the ingroup)' (Schwartz et al., 2006, p. 6). It can also impact the negotiation and alteration of cultural norms and values, therefore influencing the construction of their cultural identity (Vigil, 2004; Slootman, 2014), defined as 'sense of solidarity with the ideals of a given cultural group and to the attitudes, beliefs and behaviours manifested toward one's own (and other) cultural groups as a result of this solidarity' (Schwartz et al., 2006, p. 6). This might be especially relevant for urban youths from immigrant families who not only need to forge their personal identities but are also be confronted with 'the challenges of reconciling multiple cultural systems of reference' (Sirin & Fine, 2007, p. 152). Hence, they also need to undergo acculturation, defined as the process in which migrants and their children change and alter identities, values and cultural practices[10] (Berry, 2005).
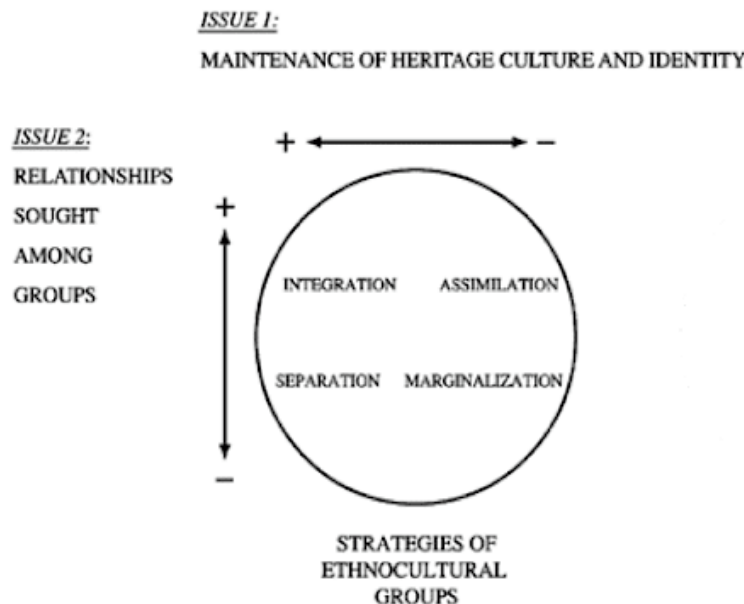

Figure 6. Berry's Original Model of the Four Acculturation Strategies (Berry, 2005))

---

[10]While this theory is often used to investigate the acculturation of migrants and their children in so-called 'new' or 'host' societies, it is important to notice that Berry (2005) explicitly states that acculturation is also a process that the members of the so-called 'receiving' societies undergo.

According to Berry (2005), the acculturation process depends on '(1) a relative presence for maintaining one's heritage culture and identity, and (2) a relative preference for having contact with and participating in the larger society along with other ethno-cultural groups' (p. 704). On the basis of these factors, Berry (2005) identifies four possible acculturation strategies (Figure 6). (1) Assimilation implies the endorsement of identities and practices of the contact group and the rejection of heritage identity and culture;(2) separation involves reaffirming the heritage culture and identity, while rejecting the contact group's cultural practices; (3) marginalization as an acculturation strategy requires the rejection of both heritage and contact group culture and identity; and finally (4) integration as a strategy forthe maintenance of heritage and contact group's identities and cultural practices. Berry (2005) notices that the acculturation strategy can be influenced ' but not determined ' by the context in which acculturation takes place. Indeed, inter-group conflict may lead an ethnic group occupying the position of the 'minority' to endorse assimilation or separation as an acculturation strategy, making the adoption of integration as a strategy more difficult (Berry, 2005).

In the Netherlands, research has tended to focus on whether, individually, Moroccan-Dutch youths identify as Dutch or not, often with contradictory results. Some scholars argue that there is a general trend of disidentification with the Netherlands that demonstrates that they are not assimilating or integrating (Maliepaard, Lubbers and Gijsberts, 2010; Maliepaard, 2012; Maliepaard, Gijsberts & Lubbers, 2012). Other scholars show that Moroccan-Dutch youths deploy different acculturation strategies and reveal that a disadvantaged socio-economic position and perceived discrimination hampers integration and assimilation (Verkuyten, Thijs & Stevens, 2012; De Vroome et al., 2014). Only a few studies have focused on the role of peers on Moroccan-Dutch youths' construction of their cultural and social identity. Quantitative research has mostly focused on the quantity of contact with co-ethnic peers, which was found to hamper assimilation and integration (Azghari et al., 2017). Similarly, other studies have focused on the negative influence of peer rejections on adopting integration as an acculturation strategy (Celeste, Meeussen Verschueren & Phalet, 2016; Lindo, 2008). In light of these previous studies,

this research tried to unveil the relation between peer groups, peer culture and the construction of Moroccan-Dutch youths' cultural and social identities.

3.3 Supportive Peers and Peer Culture

3.3.1 Peers and Peer Culture as Everyday Social Support

Urban youths from immigrant families have to deal with a variety of challenges that may lead to experiencing stress. Peers might mitigate the harmful effects of environmental stress on youths' mental health because they also can provide social support, defined as the provision of a resources by social actors intended to enhance individuals' well-being in stressful situations (Shumaker & Brownell, 1984; House, 1981; Cohen, & McKay, 1984; Cohen & Syme, 1985). More specifically, in classical social support theory, four kinds of social support are identified (House, 1981): instrumental support (e.g. providing tangible resources such as time, skills, economic support); emotional support (e.g. providing care, love, trust), appraisal (e.g. re-evaluation of the situation) and informational support (e.g. information to assist and solve a problem).

By providing some sort of support, peers and their peer culture may act as a protective factor (Stanton-Salazar & Spina, 2005; Stanton-Salazar, 2011; Juang, Ittel, Hoferichter & Gallarin, 2016) - a 'moderato[r] of risk or adversity that enhance good outcomes' (Masten, 1997, p. 6). Indeed, the support received by peers also might help the youths to cope with everyday stress, whereby coping is defined by Lazarus and Folkman (1984), as the 'ongoing cognitive and behavioural effort to manage specific (external and/or internal) demands that are appraised as taxing or exceeding the resources of the individual' (Mitrousi, Travlos, Koufia & Zyga, 2013, p. 131).

In the Netherlands, researchers have focused more on the role of community and parental support in the lives of Moroccan-Dutch youths (van Mourik, Crone, Pels & Reis, 2016; Rezai, Crul, Severiens & Keskiner, 2015; Pinkster & Fortuijn, 2009; Eijberts & Roggeband, 2016). Not much research has been conducted on peer support and on whether peer culture is oriented to social mechanisms of peer support. For this reason, this research in-

vestigated the role of peer support in the everyday lives of a group of Moroccan-Dutch youths and, more specifically, in case they share a peer culture, whether such peer culture promotes supportive mechanisms among the youth.

3.3.2 Peer and Peer Culture as Social Capital

Relationships between peers 'can prove to be an influential social currency' (Gibson, Gandara & Koyama, 2004, p.7) in navigating socio-economic, cultural and religious domains. Peers can be a source of social capital 'understood roughly as those 'connections' to individuals and to networks that can provide access to resources and forms of support that facilitate the accomplishment of goals' (Stanton-Salazar, 2004, p. 18), providing an important aspect to socio-economic mobility (Bourdieu, 1977).

As Gibson et al. (2004) and Stanton-Salazar (2004) explain, peer social capital becomes particularly important for youths from working-class and migrant families, because these youths might have limited access to sources of critical information, support, and to institutional agents, defined as 'high-status, non-kin, agents who occupy relatively high positions in the multiple dimensional stratification system, and who are well positioned to provide key forms of social and institutional support' (Stanton-Salazar, 2011, p. 1066). In this case, as Gibson et al. (2004) explain, peer social capital 'can provide access to tangible forms of support [such as] institutional resources and funds of knowledge that enable low-status students to decode the system and participate in power sharing' (Gibson et al., 2004, p. 8). Indeed, urban youths from migrant, working class families 'too can be embedded within institutional sites that generate plentiful resources' (Stanton-Salazar, 2004, p. 31). As Stanton-Salazar (2004) writes, 'despite the typically debilitating effects of their class origins, many working-class individuals do find themselves embedded in social networks where they enjoy privilege access to supportive contexts and to institutional agents and resources' (Stanton-Salazar, 2004, p. 20-21).

In the Netherlands, researchers have primarily focused on the importance of the support received from ethnic communities and parents to enhance socio-economic mobility (Crul & Doomernik, 2003; van der Veen & Wim Meijnen, 2002; Crul & Heering, 2008; van Mourik et al., 2016; Rezai et al., 2015; Pinkster & Fortuijn, 2009) and to deal with discrimination (Eijberts & Roggeband, 2016). Hence, this research aimed to investigate how Moroccan-Dutch youths' peer relationships become conduits of social capital and resources. The research also aimed to uncover whether the peer culture that Moroccan-Dutch youths share with their peers promotes supportive mechanisms that enhance the circulation of social capital and whether such social capital enables their mobility or is limited in helping their effective navigation in society.

Methods

This research relied on fieldwork conducted with 20 youths (13 males and 7 females) aged 17-26 between February and May 2018. However, data collected from the fieldwork conducted with 15 youths between January 2017 and December 2017 were also re-analysed in light of this thesis' research questions and theoretical framework. In total, 110 hours of fieldwork were conducted. Since five youths participated in both rounds of fieldwork, the total sample size amounts to 30 participants (12 females and 18 males) aged 17-26 living in Amsterdam and in Rotterdam. Furthermore, a Moroccan-Dutch teacher and youth coach, a Dutch teacher, a Moroccan-Dutch policy maker, and a Moroccan-Dutch adult were regularly consulted. In order to find participants to investigate the role and values of peers in the lives of Moroccan-Dutch youths, I relied on convenience and snowball sampling. I mostly relied on my personal network, as despite having access to Nidal's[11] organization ' the Moroccan-Dutch teacher and youth coach 'I could not immediately access participants.

Data collection relied on methodological pluralism, defined as a method that involves 'a set of traditional word-based textual methods as well as visual methods' (Katsiaficas, Futch, Fine, & Sirin, 2011, p. 123). This approach was used because visual data (e.g. maps drawn by the participants) elicit new interpretations, allow to discover hidden stories and provide new lenses for analysing

---

[11]Pseudonymous used in order to guarantee the anonymity and the confidentiality of the participant. In fact, I provided the participants with consent forms guaranteeing confidentiality and anonymity, as the research was conducted in an ethical fashion and complied with Amsterdam University College Research Ethical Standards.

textual-based data (Futch & Fine, 2014). Further-more, asking the youths to draw maps while collecting textual-based data facilitates and encourages the sharing of stories and experiences (Futch & Fine, 2014). Qualitative data specifically on the role and value of peers in the lives of the participants were collected through 12 individual and 3 paired semi-structured interviews, one focus group, participant observation, informal conversations, and peer maps[12] drawn by the participants. This data was collected in the fieldwork conducted between February and May 2018 over the span of approximately 45 hours with 20 participants (7 females and 13 males) aged 17-26.

These data sources were complemented with the data gathered between January and December 2017 in the form of participant observations: 32 individual and 2 paired semi-structured interviews with 15 youths aged 18-25 living in Amsterdam and Rotterdam.[13] Of the 15 participants, 6 drew their social support maps, while 7 also shared their life histories and drew their life history maps, as they were asked, following Futch and Fine (2014), to draw a life map that 'include people, places obstacles and opportunities on the way' (p. 46). Although last year's data collection was framed by other research questions, it contained multiple references to the role that peers have in their lives. Indeed, the life histories and life maps were especially insightful in understanding the role of peers in vital conjectures. The interviews on social support and the social support maps were fundamental to understand the role that peers provide on a daily basis.

Data reduction was done by using Microsoft Word and Express Scribe. I 'cooked' the raw field notes taken while conducting fieldwork no later than three days from the date of collection. Similarly, I transcribed most of the interviews verbatim no later than a week from the date of collection. Textual and visual data was analysed through thematic coding, which Braun and Clarke defines as 'a method for identifying, analysing and report-

ing patterns (themes) within data' (in Flick, 2014, p. 421). First, data was coded by doing open coding. The creation of code maps helped to cluster open codes into categories and then constructs, which were also described in analytical memos and reported in a preliminary codebook. This enabled me to move into the next analytical step, that of axial coding, which facilitates the understanding of patterns of similarities and differences between the narratives of the youths (Tracy, 2013; Flick, 2014). The creation of matrices for comparison of some of the youths' accounts and the writing of analytical memos aided this process. Lastly, selective coding enabled me to establish the overarching themes. The data are reported according to the emergent themes.

Two major difficulties encountered in the field work need to be addressed. The first limitation concerns the fact that I do not speak Dutch, so the research had to be conducted in English. This limited my participation in the activities that the youths share with their peers. More radically, it was problematic because nuances might have been lost in the translation from Dutch to English. Secondly, time restrictions within the fieldwork addressing specifically the role of peers in the lives of Moroccan-Dutch youths was a major limitation. First, it did not allow me to establish the kind of rapport that would have given me access to the activities that the participants do with their peers. Secondly, I could not conduct members' reflections[14], which are fundamental to identifying misinterpretations and discovering hidden biases that come from my positioning as a white Italian woman attending a prestigious college in the Netherlands.

Findings

The findings are organized in four macro sections. The first section describes the participants. The second section describes the participants' peer relationships. The third section presents the activities that the youths engage in with their peers,

---

[12]Visual representation of the relationships the youths have with their peers.

[13]Qualitative data was collected since January 2017 in the author's prior research projects: January 2017: Clothing practices of a group of Muslim women living in the Netherlands (Turcatti, 2017e); Summer 2017: The life histories of a group of Moroccan-Dutch youths (Turcatti, 2017b; Turcatti, 2018); Fall 2017: The role of social support in the lives of a group of Moroccan-Dutch youths (Turcatti, 2017d); Fall 2017: The meanings that a group of Moroccan-Dutch youths attach to the gap year. (Turcatti, Berger & Kruimer, 2017

[14]Members' reflection is a technique deployed in qualitative research whereby preliminary findings emerging from the analysis of the researcher are presented to the participants who give feedback on such analysis (Tracy, 2013). The aim of members' reflection is to make the research process more inclusive and to spot researchers' misinterpretations. (Tracy, 2013)

how these activities function as spaces for identity work, and the values, norms and meanings that the youths report to share with their close friends. The fourth and final section describes the ways in which the participants and their peers support each other.

5.1 The Participants: Similar Background, Different Life Stages

The twenty participants' were between 17 and 26 of age. Of the participants, twelve were females while eighteen were males. The majority of the participants were born and raised in Amsterdam. Only six participants were born and raised in Rotterdam. All of the participants were raised in highly diverse neighbourhoods. In some cases, the participants report to have been raised in 'a small house' along with their siblings.

Almost all the participants have both Moroccan parents. In many cases, the participants' fathers came to the Netherlands as guest workers in their twenties. In the cases in which the participants' parents met in Morocco, the participants' mothers reunited with their husbands in the Netherlands at a later stage. There are also instances in which the participants' parents met while already living in the Netherlands.

At the time of the research, the majority of the participants were still studying. The majority of the participants who were still in education were attending MBO (eleven males and four females), four young women were enrolled in HBO, Yassine was about to graduate from VMBO, while Sarah was about to graduate from VWO, three young males were doing their bachelor studies, and Lamia was about to obtain her master degree.[15]

At the time of the research, seven of the participants were not studying. Lina was not studying because she had just graduated from VWO and

she was taking a gap year to save some money. Zakaria and Said were not studying because they obtained their Bachelor and HBO degrees, respectively. Similarly, Fatima was not studying because she had graduated from her Master studies. Jason, Ali, and Rachid were not studying because, driven by different motivations, they decided to work instead.[16]

Independently of whether the participants were still studying or not, the majority of the youths that participated in the research is working. The participants who study also hold part-time jobs at grocery stores or shops. This is because they want to make some pocket money and help their parents with the familes' finances. The participants who are not studying hold full-time jobs.[17]

5.2 The Youths' Peers and Peer Relationships in Everyday Life

This section describes (1) who the peers with whom the participants interact are; (2) the way the youths classify their peer relationships in terms of strength; (3) and the explanations that the youths provide when asked to account for who they befriend.

5.2.1 Peers and Context of Interaction

The youths often classify their peers according to the context in which they interact (Figure 9). In terms of gender, the peers with whom they interact on a daily basis are both females and males. The educational level, the background, and religion of the peers with whom the youths interact also change according to the context of interaction. The peers from the participants' neighbourhoods tend to be 'diverse' or 'multicultural' in terms of background rather than Dutch. This is because almost

---

[15]VMBO (*Voorbereidend Middelbaar Beroepsonderwijs*) refers the vocational-training based sector of the Dutch secondary school system. Upon completion of VMBO, students can enroll in MBO (*Middelbar Beroepsonderwijs,*) higher-level vocational training, after which participants can enroll in HBO (*Hoger Beroepsonderwijs*) or enter the labor market. HBO is the highest level of the Dutch vocational training system. After HBO students can either enter a program or enter the labor market. VWO (*Voorbereidend Wetenschappelijk Onderwijs*) is the highest level of secondary education in the Netherlands, and completion of VWO allows the student access to Dutch universities.

[16]Jason and Ali quit their studies at HBO, while Rachid quit his studies at MBO. While Ali decided to quit HBO because he realized that 'studying was not my thing',this was not the case for Rachid and Jason. Rachid quit his studies because he realized that it would take him too much time before he could start working, save money and form a family. Jason did not finish his HBO studies because in the school 'everyone was racist'.

[17]Fatima, Said and Zakaria hold jobs that require the skills they acquired in higher education: Fatima works as a psychologist with young refugees, Said is an engineer, while Zakaria is an assistant professor at the university where he studied. Although Jason quit his HBO studies, he was able to take advantage of his knowledge of Dutch, English and Arabic and to work at a translation company. Lastly, Ali works as a security guard, while Rachid works in his family's restaurant.

all of the participants grew up in diverse neighbourhoods of Rotterdam and Amsterdam with a low share of Dutch people. Furthermore, in their neighbourhood, the youths perceive the average educational level lower to that of 'Dutch'[18] or 'white' neighbourhoods.
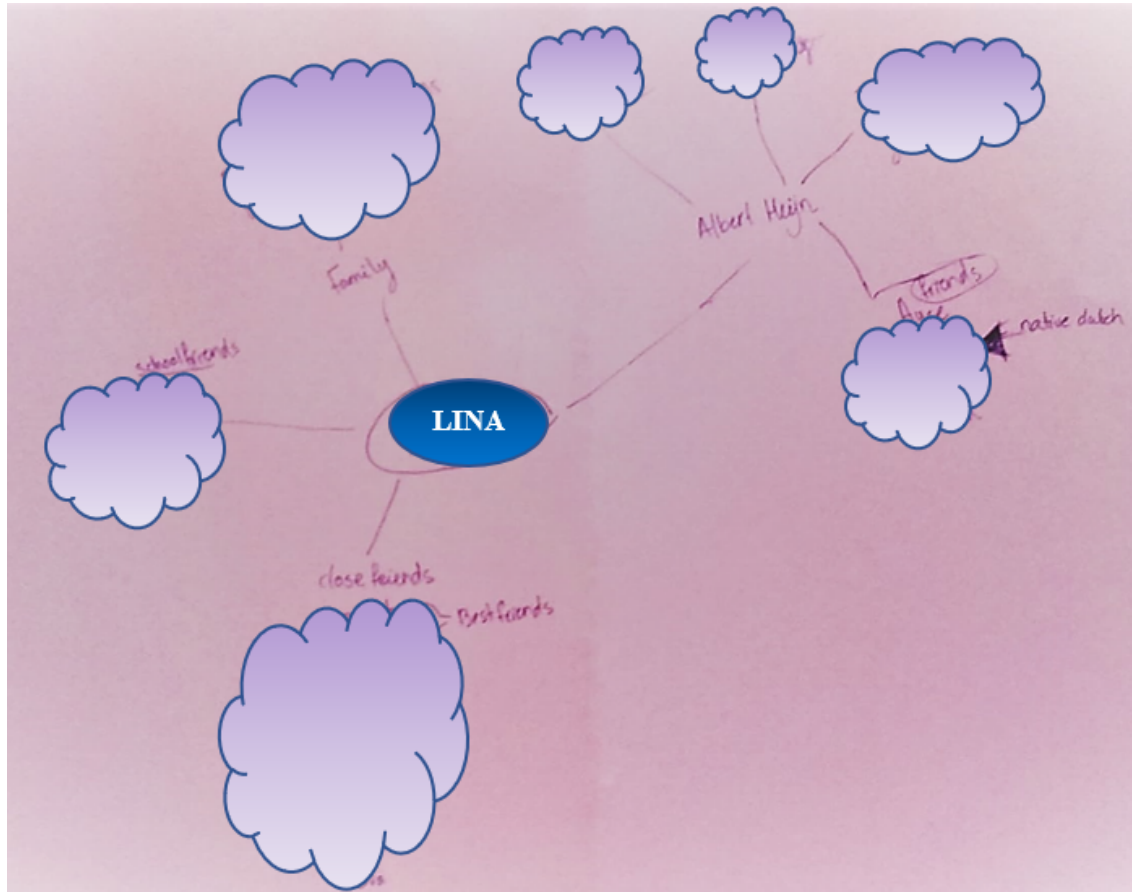
Lina's[19] Peer Map. In her peer map, Lina classifies the youths with whom she interacts according to the context where she meets them (e.g. school, family, work)

The youths also explain that, depending on the high school that they attend(ed), the educational levels, background and religion of school peers vary. The participants who attend(ed) VMBO, MBO or HAVO do not come in contact with Dutch peers often. Furthermore, in these schools, the youths report to be surrounded by many Muslim peers. On the contrary, those who went to VWO or Gymnasium tended to be more in contact with Dutch and non-Muslim peers.

[18]It is important to highlight the fact that the youths use the nationality of their parents to identify their own background and that of their peers. It is necessary to specify, however, that this terminology does not imply that the youths do not identify as Dutch. Indeed, when asked how they identify themselves, the youths tend to respond 'both', meaning Moroccan and Dutch. The youths explain that the reasons why they use this terminology is to honour their background and also because this is how they are often addressed. The youths explain how Dutch people often do not accept the answer 'Dutch' to the question 'Where are you from".

[19]The actual name of Lina and of her peers have been shaded in order to protect the participant's and her friends' anonymity. The same was done for all of the maps drawn by the youths.

Lina's[20] Peer Map. In her peer map, Lina classifies the youths with whom she interacts according to the context where she meets them (e.g. school, family, work)

Similarly, depending on the level and kind of tertiary education attended, the background, and religion of university peers vary. The participants attending HBO or the Vrije Universiteit (VU) in Amsterdam report to be in contact with 'multicultural' and 'diverse' students. Zakaria, who is the only participant who studied at Universiteit van Amsterdam (UvA), reports how his university peers 'are mostly Dutch, because the UvA is over-represented by Dutch people' (Figure 8).

In general, the youths tend to work in environments that feature Dutch and non-Dutch colleagues who can be of their age or older. Only Aya and Rachid report to work mostly with Moroccans, while Lamia works in a 'white' environment. Some of the youths reported that work is often the domain where they met Dutch peers for the first time.

5.2.2 Strength of Peer Ties: From Weak Relationships to Friendship

The relationships that the participants establish with their peers vary in strength. Some are weaker in the sense that they are limited to the context where the youths interact (e.g. school or work). With other peers, the participants develop stronger ties. These are the peers that the participants start to address as 'friends'. This, however, does not mean that they meet often (Figure 9). The participants develop even stronger relationships with some of their peers. Finally, they all identify and elect the peer with whom they have the strongest

tie (the 'best friend', the 'closest friend'). Figure 10 summarizes the types of relationships that the participants develop with their friends in terms of strength.

The youths' close friends tend to be of the same age, gender and educational level as the participants. In terms of diversity, only few participants tend to have close friends that are Dutch, 'multicultural'[21], Muslims and non-Muslims. The majority tend to have Muslim and non-Muslim close friends but with a background other than Dutch. Other participants tend to have close friends with a background other than Dutch but who tend to be all Muslims. On the contrary, Anas' close friends are all non-Muslims but of diverse background. Rachid is the only participant who reports that his close friends are only Moroccans.

5.2.3 Emic Explanations of Strong Peer Ties

To the question "Why do you consider these people your close friends?", the participants provide multiple explanations which can be categorized in four different categories: (1) exposure to contact; (2) geographical proximity; (3) 'being like me"; and (4) outcomes of the relationship.

Exposure to Contact

Most of the participants report to have developed close friendships with the youths with whom they have engaged in prolonged contact in contexts such as school, work or the activities done

---

[20]The actual name of Lina and of her peers have been shaded in order to protect the participant's and her friends' anonymity. The same was done for all of the maps drawn by the youths.

[21]Term used to identify peers whose parents are not Dutch.

in the neighbourhood. Zakaria explains that being exposed to contact enabled them to get to know each other and develop shared memories, which contributed to strengthen their bond:

> "We have been through a lot together, fun things, things we are still talking about it, laughing about it, stupid things that we did in high school...that connects you as human beings. Memories, you share the same memories." (Zakaria, 22, Amsterdam)

However, the majority of the participants have built close friendships while being in high school, which seems to explain why the participants' close friends have similar educational levels. It also partially accounts for the diversity of the participants' close friends. It was shown earlier how the background of the peers with whom the participants come in contact is shaped by the schools' characteristics. Indeed, when the youths already live in a 'multicultural' neighbourhood and attend a 'multicultural' school, it becomes harder to befriend Dutch peers. The contrary is also possible.

However, Ali, Rachid and Yassine, who attended VMBO and MBO, report how they do not have friends from high school. According to Ali, if he had maintained his high school friendships, he would be 'in jail or shot dead' by now. Furthermore, it is important to notice that not all the participants had a stable educational path and were therefore able to develop long-lasting relationships with their schoolmates. Rachid, for example, was downgraded from VWO to HAVO after the first year. However, at HAVO the teachers failed him once more and told him 'you have to change schools', 'you're a bad student'. Further downgraded to VMBO at the end of the year, he had to look for another school.

Geographical Proximity

Living close to each other also seems to be a condition that enables the development of close friendships. Geographical proximity seems to increase the chances of spending time together outside the contexts of interaction. Furthermore, some male participants report how they have befriended their peer neighbours, even if they did not go to school together.

"Being Like Me"

Furthermore, the youths seem to identify 'being like me' as an important factor for the development of close friendships. Most of the participants explain that they tend to befriend those with whom they share similar personal interests. In addition, some of the highly educated participants tend to specify that they find it easier to 'connect' with those who have the same educational level. Similarly, the fact that they tend to befriend peers of the same gender might also be understood in terms of homophily.

More fundamentally, some of the participants report to find it easier to befriend those who are 'also Moroccans'. This, however, does not mean that there is one stable, essential concept of what it means to be Moroccan as it is a concept constantly in flux upon which the youths might not always agree. Rather, befriending 'Moroccans' is about imagining to share a set of Moroccan cultural values which, in reality, might or might not be the case. Zakaria's words seem to be showing precisely this imagination:

> "...you share in a certain way the same background, the same principles, but you can bond with them because they are from your own [...] you share the same culture so you connect with them more easily than you connect with the Dutch guys." (Zakaria, 22, Amsterdam)

Other female participants stress how it is easier to bond with the peers with whom they share the same religion[22], rather than background. In Lina's words:

> "I feel more similarities for example a Suriname-Dutch girl [...] or a Turkish-Dutch girl, because we all have the same problems ['] The point is ['] we are far more behind than the natives, because the natives grow up here, they study here and most of their parents have a good job, they live good [...]

---

[22]Even in this case, finding it easier to bond with the peers who are also Muslims does not mean that the participants and their friends have the same understanding of what it means to be Muslim or how to be Muslim. This is highlighted by the fact that the participants report how sometimes their religious practices differ from those of their close friends.

> But most of the people that live in the Netherlands and are coloured...they have either their parents or their grandparents that came here to work, didn't have much money, they were here for a reason, and their kids also grew up in a poor family, which turns out they didn't have luxury to study and all those things." (Lamia, 23, Almere)

To this, Aya adds that she connects more easily with the sons and daughters of migrants also because of the fact that they might share a feeling of not being fully accepted by the native Dutch peers. Najat explains that, besides the feeling of not being fully accepted because of their background, she finds it easier to connect with them also because they all share the struggle of making sense of who they are:

> "[..] it is a typical thing of this second generation of migrants [..]When we are here in Holland people view us as 'Moroccans, foreigners' and when we visit Morocco they see us as 'Dutch people'['] Friends with a migration background share this struggle also." (Najat, 18, Amsterdam)

Yet, some of the participants explain how they might bond with Dutch peers provided that they also confront socioeconomic challenges and/or they are used to diversity in the sense that they have been raised in diverse contexts. Omar specifies that he strongly connected with Lasse, his Dutch school mate, precisely because of having lived similar 'experiences, struggles'. In Omar's words:

> "Because he [Omar's Turkish friend] shares so many things... ['] he knows a lot about life ['] Experiences, struggles, good things ['] it's just environment [...] Also, background...because he's from Rotterdam South which is not the best place to grow up. That's where he came from." (Omar, 21, Rotterdam)

Interestingly, Omar and his Turkish friend explain that they 'don't see Lasse as a Dutch [person]'. Omar and his friend seem to associate 'being Dutch' with not having to face any socio-economic

challenge and economic privilege, which they associate with 'being white'. Similarly, the participants who report to connect well with the Dutch youths who are 'used to diversity' do not see them as 'regular Dutch' peers, insofar they construct the 'regular Dutch' person as the one that is not used to diversity and that contributes to the stereotyping of others. For Lamia, this is what makes the 'regular' Dutch 'white'.

Outcomes of Peer Relationships

Lastly, many of the participants stress the fact that the way the relationship with their peers is experienced is important. Some male participants emphasize that in order to strengthen the bond with their peers, it is important that their peers are trustworthy and loyal. Furthermore, in order to build a closer relationship with their peers, most of the youths emphasize the importance of not feeling judged. As Najat explains:

> "Well, the most important thing is that I feel good when I am with them, that I can be myself with them, because I had period that others would bully me, but at secondary school who took me as who I am ['] that's important that you can be yourself and not be judged." (Najat, 18, Amsterdam)

The expression 'I can be myself' seems to indicate the importance of not being 'judged' by other Muslims for how they practice their religion. It also means not being judged by other Moroccan youths for not being 'fully' Moroccan ' as in the case of Najat and Soufiane who only have one Moroccan parent ' or for what they do and with whom they do it. Omar's experience openly addresses this:

> "I heard so many times that I am so Dutch and I don't know...they can't tell me why I am Dutch ['] they tell me...you have Dutch friends, you go to Dutch parties...just dumb stuff...." (Omar, 21, Rotterdam)

It is important to notice that not feeling judged becomes an important factor for establishing relationships with Dutch peers who are perceived as 'white' either because of their wealth or because of the fact that they have never been in contact with

people of a diverse background. Said powerfully summarized the conditions under which the participants can get along with Dutch peers:

> "I can get along with both [Moroccan and Dutch] as long as they don't judge me. If they talk with me like I am a person, then we can talk. But if they talk to me like all Moroccans are criminal [...] The moment you judge me, I am out. I am not gonna fight for who I am." (Said, 26, Amsterdam)

The importance of not being judged because of who they are, the way they express their religiosity, the way they are Moroccans and that the very fact of being Moroccan seems to originate from previous negative experiences where they felt discriminated or not accepted.

5.3 Peers as Spaces for Identity and Meaning Making

This section reports (1) the diverse activities in which the participants engage with their close friends and (2) how these activities become spaces for reflecting on who they are. It also presents (3) the meanings, values, and norms that the youths indicate to share with their close friends and (4) how such shared meanings are the foundation on which they define their place in Dutch society.

5.3.1 'More Than Just Fun'

The activities in which the participants engage with their friends seem to be influenced by the youths' personal interests and by what they enjoy doing. For instance, Anas plays videogames with his friends at each other's houses, while Aya and Zineb like to have lunch together. The activities done with their friends are also constrained by their everyday life commitments. For instance, Soufiane does not go clubbing when he is preparing for his university exams. In addition, the activities that are done with the peers seem also to be shaped by the religious and cultural values endorsed by the participants. Lastly, the female participants point to the fact that the activities they engage in with their peers are also shaped by parental supervision which limit the extent they attend parties or prevents them from hanging out 'in the street'.

Despite variation in the kinds of activities that the youths enjoy with their close peers, all the participants report to highly value the 'deep conversations' that they have with their close friends. For the majority of the participants, 'shopping together', 'sitting in the car', 'chilling', 'smoking shisha' are all occasions to 'talk about life'. Omar powerfully expresses this:

> "...Whenever my friends are just talking about life all day every day. We smoke a lot of shisha and we are very deep persons so every day we have a question about life [...] Every day we tell each other our life stories with my friend." (Omar 20 Rotterdam)

What is evident from Omar's quote is that close friends offer a space for reflecting on the everyday life issues that the youths confront, on who they are and who they want to be. Hence, 'talking about life' seems to become a coded expression for identity work and for meaning making.

5.3.2 'Chilling' as Identity Work

This subsection presents how the activities and conversations that the youths have with their close friends seem to offer a space to work out (1) their personal identities; (2) their Muslim identity; (3) their cultural identity as Moroccans and Dutch; and (4) their social identity in terms of social distance from who the youths perceive as 'different'. With their close friends, the participants often report discussing educational and job-concerning choices. Furthermore, the participants seem to discuss and reflect upon the opportunities that they are offered and how to manage them. The joke that Aya recounts might be an example of this:

> "For example, she [describing a friend] works at the government. She worked there and she hated it [...]. And I told her, like, if that place was the only opportunity, how long would you drop out of school' And she was like, I like that way of thinking! [...] She was like laughing about it and she wasn't annoyed that I said that [laugh] She hated it so much!" (Aya, 21, Rotterdam)

Close friends also seem to offer a space to reflect upon the different ways in which the youths

live their religiosity. Najat's words seem to point to this process:

> "My humour tends to be a bit dark, about taboo subjects that for some people might be too much, luckily my friends can appreciate it. We all have different lifestyles and some of us are religious and some not, so often we talk about that subject." (Najat, 18, Amsterdam)

Almost all of the participants seem to indicate how close friends are also spaces for comparing and evaluating what they perceive to be the cultural differences between Moroccan and Dutch culture. For example, Aya reflects with her Moroccan friend on what she perceives to be the place that sexuality occupies in both cultures:

> "You know some things are not normal for you. Some Dutch are so open about stuff and we are like woooooo [laugh] How do I say that' Our culture is like very taboo kind-ish and when here people talk about stuff, sex and stuff, whoooo [...] We don't even say that word. It's not that we make jokes about being open, but [about the fact] that it's no usual for us to say that at home." (Aya, 21, Rotterdam)

It seems to be the case that close friends tend to embrace similar acculturation strategies.[23] When asking Moroccan participants belonging to the same friend group how they identify themselves, they often provide similar answers. The majority of the participants report that they identify as Dutch and Moroccan. Lamia's response to the question 'how do you identify yourself" is representative of most of the answers the participants provided:

> "We have the in between the Moroccan and Dutch, a way that suits both, you know' [...] I am like...in between. I am not Dutch, I am not Moroccan either but I think...I think I am a Dutch-Moroccan, because I feel like the Moroccan people

who live in the Netherlands [...] they have their own culture, they have their own way of acting with people." (Lamia, 23, Amsterdam)

There are exceptions, however. Yassine and his close friends' response to the same question is strikingly different. Yassine calls himself Moroccan because he feels Moroccan, not just because his parents are Moroccans. During summer 2017, I also asked the same question to his Indonesian and Surinamese friend, while having dinner. The response was strikingly similar. However, Yassine made an important specification:

> "Maybe I want to feel Dutch [...] but they didn't give me the chance! [...] They always put us outside, they are always bad in the news, always bad people, bad thing, bad Muslim they did that and that..." (Yassine 20, Amsterdam)

Yassine's and his friends' lack of identification as Dutch seem to derive from the perception of having been negated the same socio-economic opportunities and from their everyday experiences of discrimination. Rachid's story might exemplify this feeling:

> "Maybe how I walk' how I dress' No seriously. They look at me and they think this Moroccan is crazy. If sometimes I walk behind Dutch people [...] they keep, keep, keep watching behind them. What do they think? That I am going behind them to stab them in the back? Pathetic! It's sad! Sad people, you know." (Rachid, 21, Amsterdam)

Rachid's quote wants to point to the fact that the participants also discuss their experiences of inequality and discrimination with their close friends. Lina and Sarah's conversation might be seen precisely as an example of this, while also wants to show how females and males' experiences of discrimination seem to be qualitatively different:

> Lina: "My friends have the same opinion as me that we are treated unfairly sometimes because of the way a certain group of people..."

---

[23]It is nevertheless important to specify that the Moroccan youths who are in the same group of close friends do not necessarily retain or reject the same Moroccan values and practices as much as they do not reject or embrace the same Dutch values and cultural practices.

Sarah: "I don't get it that often."
Lina: "Because we are girls and we don't wear the hijab."
Sarah: "Boys get more bad attention..."
Lina: "With girls they think that we are oppressed, especially if you wear the scarf, and with guys they are...oh yeah, you are nasty..." (Lina, 19, and Sarah, 19, Amsterdam)

It is important to specify that the youths also address in their friends' group the way other Muslim and Moroccan peers and adults might judge them for the way they conduct their lives. For example:

Sarah: "[...] you also have the very strict ones. It's not the largest group..."
Lina: "[...] the really religious. They wear the hijab and they do it for themselves but...you also have the girls that for this they don't go to the Christmas parties..."
Sarah: "[...] you also have those who wear a headscarf and they judge other people who don't." (Lina, 19, and Sarah, 19, Amsterdam)

By doing so, the youths seem to negotiate the social distance to maintain towards those who they perceive as 'different' from them and from those who have treated them 'unfairly'.

5.3.3 Sharing the Same 'Mindset', the Same 'Culture'

As Soufiane says: "we're kinda creating our own little thing, as we weren't accepted in the first place." In fact, not only do the participants report to 'talk about life' with their peers, they also explicitly state to share 'the same mindset' or 'the same mentality'. Others use the term 'culture':

"I feel like the Moroccan people who live in the Netherlands [...] have their own culture, they have their own way of acting with people. We are living with a lot of different backgrounds, cultures' together, being friends, hanging out [...] I feel more similarities, for example, with a Suriname-Dutch girl who is originally Surinamese but she lives in Holland, or a Turkish-Dutch girl, because we all have

the same problems, we deal with it in the same way and we have our way of getting along with each other that defines our culture." (Lamia, 23)

This section presents what seems to be the dimensions of such a shared 'mindset' and 'culture'.

"Make something with your life!"

All of the participants report that in their close friends' group, they share the same 'drive to make something with their life'. However, there seem to be differences between female and male friends' groups on what it means to 'make the most of it'. Most of the male participants, irrespective of their educational level, report that for them and their friends to 'make something with your life' means becoming ' in Zakaria's words - 'wealthy but healthy', which translates in becoming rich through legitimate means. For some of the participants, the desire to 'make money' seems to reflect the desire to achieve a better socio-economic position than their parents and conduct a more 'comfortable' life. Omar's words address this desire:

"I hope in four or five years to end my bachelor with success...after that...the really serious things come...master, job, a house [...] my house has to be nice and big so I feel comfortable when I am in it. All my life I have been in some tiny houses. If you live with six people in a house with two rooms, you don't feel comfortable." (Omar, 21, Rotterdam)

Although the female participants also report to have been raised in a 'small house' by parents who did manual jobs, it seems that for them and their close friends 'making the most of it' means achieving a high-status job. For some of the female participants, this desire seems to arise from the comparison that they draw with the women in their families and the imagined comparison with the women who live in Morocco.

It is important to notice that in the friends' groups composed of youths who have attended HAVO and VWO and/or are attending tertiary education, 'making the most of it' means 'following your dreams', 'making money' by doing what they like. The youths report that they share this with

their peers because they were all given the educational opportunities that not only allowed them to discover what they like, but that will open the doors for them in the future. However, also in the friends' groups composed of youths who have attended or are attended VMBO or MBO there is the desire to 'become someone' by doing what they like. Some of them want to become sports teachers, Anas wants to found his own media company, others want to continue their studies and proceed to HBO.

There are nevertheless exceptions among those enrolled in lower levels of education. Yassine and his Indonesian and Surinamese friend - who are currently studying informatics at VMBO ' feel they still have not figured out what they like, who they want to become and the way to 'make money'. Indeed, they do not like what they are studying and they report to be 'wasting' their time at school as they perceive that the informatics studies that they are currently doing are only for 'stupid' people. In Yassine's words:

> "I don't see something in school, I can do much more. [...] I can build my own company, I can work, I can do everything..." (Yassine, 20 Amsterdam)

Yet, although Yassine reports that he 'can do some much more', to the questions what do you want to do' and what ideas do you have', he keeps responding 'I don't know'. Furthermore, in other instances, some of the participants did know who they wanted to become, but it seemed that they have given up on that dream. Rachid wanted to become a lawyer, but by being downgraded in education, he realized that it would have taken him too much time to 'get there' through education and eventually decided to work at his family's restaurant in order to start saving money to build a family in few years.

All of participants are aware of the fact that they develop this feeling and desire to 'make something with their lives' in a context that expects little from them, which is precisely what, in some cases, the participants want counteract. The youths, irrespectively of education, gender and opportunity structure, seem to be eager to show that they are better than the way they are portrayed in the media and perceived by some of their co-citizens. The way Najat objects against his school being labelled as a 'black school' might be seen as an example of this feeling:

> "I didn't really like [the term 'black school', used in Dutch slang to refer to schools with many immigrants in the student body] because [...] We did well! The only difference from a black school and a white school is that we came from, no, our parents came from different countries!" (Najat, 18, Amsterdam)

Najat also adds that her desire to 'make the most of it' originates from the drive to oppose the stereotypical ways in which Muslims are perceived nowadays:

> "I try to be a good person, I just try to live my life, I go to school, I work, I try to be healthy somewhat, just do the basic things [...] set a good example because I am wearing a scarf and I am showing all the people that I am Muslim and I represent them in a way and I have to show that not all Muslims are bad..." (Najat, 18, Amsterdam)

More specifically, the male participants want to show others that they are not the '(stereo)typical Moroccan guy', which Ali describes in the following way:

> "He will wear a sport suit of a football club, a soccer team. The everyday Moroccan wouldn't reach the success like Said have. Everyday Moroccan the first thing when he is 12 or 13 and you ask him what you wanna do with your life, he will tell you he wants to be a drug dealer...an everyday Moroccan...Now I am talking only about the males. Fast cars and stuff like that, but they don't want to work for it in a legit way." (Ali, 24, Amsterdam)

The female participants, in particular those who wear the hijab, explain that they desire to 'become someone' to show the world that Muslim women are not 'oppressed'. In Najat's words:

> "It's for us to show them not that all Muslims are terrorist, close-minded that

we can't go outside because our father doesn't want to, but that we are free, we are also modern, we also study, we also want a good life." (Najat, 18, Amsterdam)

'Help Each Other Out'

Almost all of the participants report how among their close friends' group they share the idea of 'helping each other out' in their everyday life. Even if the participants do not mention it explicitly, they often provide examples of how they help each other within the friends' group or clique. Aya's words seem to point precisely to the value and norm of 'helping each other out' as an organizing principle of the friendships she establishes with her peers:

'We are gonna help each other out, and they [referring to her fellow Dutch fellow students] are like 'no, I have to get there on my own'; that's different." (Aya, 21, Rotterdam)

Some youths explain that helping each other out through their networks and skills become important insofar it will help them with their educational and work achievements. Ali explains why he thinks, for instance, that his network of friends is 'gold':

"If I apply for a job, I don't care what it is. But you won't have a shot if you see six other white dudes. [...] My niece got her master and it took her a year and a half to find a job! and she graduated with Lauda! If she's havin' a hard time to find a job where there is job opportunity, what am I supposed to do?" (Ali, 24, Amsterdam)

'Be Used to Diversity'

Almost all of the participants report that all of their friends are and need to be 'used to diversity'. Hence, being used to diversity seems a value and a norm to comply with while engaging in the friends' groups. This can be deduced by the function of 'cohesion' that 'being used to diversity' has in Soufiane's friends' group, which he describes as a 'community'. Being used to diversity means, in Sofiane's words, to 'accept people as who they

are, acknowledge differences and try to learn about them'. For Soufiane, for example, it means that non-Muslim friends make an effort to learn about the participants' religious practices:

"It means that you accept people for who they are, it means that you have at least a bit of knowledge from other people's culture [...] You're up to date, you're interested in other people's culture and that's important for cohesion in our community." (Soufiane, 23, Amsterdam)

Aya's response to the questions *'What have you taught to your friends and what have you learnt from them"* seems to point to the fact that also the participants make an effort to learn about the cultural and religious practices of the friends who are not Muslims or Moroccans:

"I learnt a lot about Colombian culture, the slang, how Colombia is, because [...] when you go to Colombia and always people think it's dangerous but it isn't dangerous if you watch out. And she is Caribbean and she taught me that Caribbean don't necessarily break stuff...He taught me that the Moroccan and the Iraqis cultures look like each other..." (Aya, 21, Rotterdam).

Most importantly, being 'used to diversity' is also about respecting and trying to understand the differences upon which friends might not agree, instead of pressuring for assimilation of ideas and practices. Soufiane provides an illuminating example of this:

"...yeah, because everybody you see here, except maybe for this guy, see the other people often enough to understand the other person. ['] That's always differences. For example, about women [...]. if my Dutch friends would come across their sisters very drunk hooking up with a guy, they would be ok with it. But my Moroccan friends would dislike it, of course!' (Soufiane, 23, Amsterdam).

Some of the participants explain that this value of being used to diversity originates from the fact

that the participants and their friends have been raised among different cultures and have learnt how to be used to diversity:

> "These people you see here, most of them, grew up in a diverse culture, in Amsterdam South East, or in Amsterdam East... So, they are used to be around people from Suriname, from Curacao, from Morocco. So, they are used from all of these different cultures" (Soufiane, 23, Amsterdam)

However, the importance of complying with 'being used to diversity' as a value and norm within the friends' group seems also to arise from the participants' and their friends' experiences of not feeling accepted for who they are in other instances. Zakaria's words seem to exemplify this:

> "...these people [referring to his Dutch university peers] make me feel very Moroccan again, the university friends. But the high school friends make me feel like it doesn't really matter what you are. And then these friends [referring to his Moroccan friends] make me feel like I am a kinda of a Dutch guy again...They make me feel like I am maybe more Dutch than them, and these friends with the multicultural aspect [referring to his close friends from high school]. Who you are doesn't really matter cause' you know" (Zakaria, 22, Amsterdam)

'Be Open Minded'

Lastly, most of the participants report that all of their friends are and need to be 'open minded'. Being open minded is about not judging their friends, but more in general, 'people' who are 'different' because of their religion, sexuality, background, nationality or actions. Najat explains that to 'be open minded' is to see those who are 'different' as 'humans', because in Najat's words 'we are all humans'. Lina provides a concrete example of what it means to be open minded:

> "I need to have open minded friends. I have a lot of open minded friends, I don't like close minded friends. And a lot of Moroccan or immigrant people are

a bit close minded because of the way they grow up. [...] that are not open to new friends or guys or...they don't like homosexuality. [...] So, I want open minded friends I don't want the type of people that judge others just because they are gay. Do you understand what I mean?" (Lina, 19, Amsterdam)

For the participants and their friends, it seems that being open minded is also about not preventing themselves from engaging with those who are perceived as 'different'. Najat clearly explains this with regards to religious differences:

> "...Because I am Muslim, sometimes I do different things, I hang out differently, but that's not something that's up to you, as a child you are smaller and then you look at those different things. You think 'oh no, he's a native Dutch and I cannot hang out with him, because he does this, he celebrates this, he eats this' [...] when I got older I got more aware of the fact that I can still hang out with people that are different than me in no matter what." (Najat, 18, Amsterdam).

Zakaria explains that he and his friends are open minded and that they think that 'nationality is bullshit' because they grew up with diverse people, which enabled them to see beyond social boundaries:

> "I grew up with Dutch people around me, Surinamese people around me, Turkish people around me [...] I have been friends with those people from elementary school and [...] it's not a hypothesis, it's like a law - if they grow up together they will be less discriminative, less judgmental in any kind of ways!" (Zakaria, 22, Amsterdam)

Some participants specify that having positive contact with Dutch peers made them discover that Dutch 'kids' can also be 'cool kids'. However, also the participants who report not having close Dutch friends or not having been in contact with many Dutch peers share this mentality with their friends. Furthermore, even those who have only had negative experiences with Dutch people and perceive to have been denied opportunities, conclude that

whether people are 'good' or 'bad' depends on the person, rather than on their nationality, culture and religion. This awareness and conclusion seem to originate precisely from their own experiences of being associated and categorized as 'criminals' or 'terrorists' because of their background and religion. Rachid's words seem to indicate this:

> "Dutch people' [...] Some of them you know...You have racist people. But not all of them, you also have good people, Dutch people that don't care [meaning that they are not prejudiced] about Moroccans and they see everyone as the same [...] Everyone has the good and the bad people, everyone has the people that steal, the good and the bad" (Rachid, 21, Amsterdam)

5.3.4 From Peer Culture to A New Way of Being

Some of the youths explicitly indicate that these shared values and meanings make them unique. In particular, embracing the values of 'being used to diversity' and 'being open minded' is what makes them different from the peers who grew up in a 'single' culture or refuse to interact with those who are 'different'. The fact that other participants do not report this explicitly does not mean that they do not see themselves as unique. It might also be the case that they are not aware of this yet or that they do not have the vocabulary to express this. Most of the youths, therefore, report to be different from the Dutch who grew up in small villages, who are only accustomed to Dutch culture and base what they know about Moroccans on what they see in the media:

> "In exchange in Lisbon I met 4 guys from the Netherlands which I used to hang out a lot there; from Nijmegen, Utrecht, Groningen, places like that. They are completely different in everything! [...] None of them knew any Moroccan guys [...] they don't know reality, they only know what's in the media". (Soufiane 23 Amsterdam).

The majority of the participants also report to being different from the 'immigrants or the 'Moroccans' who refuse to interact with Dutch peers or anyone who is different from them or that refuse to go to Christmas parties because they are Muslims. This feeling is best represented by Anas's words:

> "There are a lot of Moroccans, Dutch Moroccan people who really also still look at Dutch people with disdain, they say 'they don't trust us, they will never accept us, I will stick to the Moroccan way, I only will hang out with my circle of Moroccan people in the neighbourhood" [...] but the thing is that they are not giving Dutch people a chance!.' (Anas 19 Amsterdam).

It is fundamental to notice that Soufiane explains that to embrace this way of being, the parents' nationality does not matter. Soufiane reports that even his Dutch friends would rather identify as Amsterdammers[24]

> "They [referring to his Dutch friends] also know our story and they have been around us, and they realize how short minded many people are in the Netherlands. So, they kinda grasp the whole picture and how Amsterdammers are ahead of their times..." (Soufiane, 23, Amsterdam)

This last quote shows how some of the participants see themselves and their friends 'ahead of their times' precisely for their own way of being. In Zakaria's words, "c'mon dude, this is the twenty-first century!".

5.4 Peers as a Valuable Resource

For Moroccan-Dutch youths, peer also become a valuable resource. This section will present how close friends contribute to each other's psychological and emotional well-being and come to share in some cases their skills, knowledge and resources.

---

[24]Indeed, it is important to notice that, while most of the participants who report being 'different' because of the values that they share with their friends do not give it a name to this 'way of being', some of the participants do so. Said and Zakaria label this way of being as being 'multicultural'. Ali uses the expression 'citizens of the world'. Jason describes it as being from 'Rotterdam' or even from his 'hood'. Soufiane labels himself as an 'Amsterdammer'

5.4.1 Peers for Psychological and Emotional Wellbeing

Close friends seem to positively contribute to the participants' psychological and emotional wellbeing by virtue of being 'always there'. The participants seem to evaluate the fact that they have close friends who are 'always there' as a positive contribution to their lives, insofar they do not feel lonely and always have someone to count on. Most of the participants report explicitly or through examples how their close friends are an important source of emotional support in emotionally challenging moments. The participants explain how their close friends become a safe space where the participants can relieve their emotions without being judged. Furthermore, some participants report how their close friends play a function of 'distraction' that they evaluate as beneficial as it enables them to diverge their attention from sources of stress. For instance, Zakaria highlights how his peers allow him a space where he can distract himself from everyday commitments. Other participants emphasize the importance of their close friends to diverge their attention from emotionally and psychologically challenging circumstances. Aya's words precisely express this function:

> "When I meet up with [my friend] Zineb, I just come for fun, because if at the moment the home situation sucks, you just want to go outside and have fun..." (Aya, 21, Rotterdam)

Some of the participants also indicate how their close friends contribute to make the participants more confident with their encouragement. For this reason, Aya describes her friends as 'cheerleaders'. Lina provides an example:

> "Last year it was my exam year and I was a bit down last year because my mum [...] would tell me all the time you are not going to make it, you have to work harder. So... I talked about it with my friends and they just told me that...because my mum...she didn't really believe in myself and my friends told me you are going to make it, you

have to believe in yourself, you are going to pass your exams...." (Lina, 19, Amsterdam).

The participants and their peers also seem to help each other to re-evaluate stressful circumstances. Youssef seems to precisely point to this process when saying that his friends 'make me think about the situation from another perspective'. Soufiane provides a specific example:

> "My girlfriend taught me...that I don't need to do stuff but that I want to do stuff. Moet niet, je wil.. So, when I am overthinking in bed about all the stuff I need to start doing or...she always says, 'no you don't need to, you want to". (Soufiane, 23, Amsterdam)

Lastly, some of the participants explicitly address how their close friends offer a space to 'fit in' when the participants perceive they are not fully accepted by 'society', or when they perceive, as Aya, to 'be born in the wrong country'. Soufiane expresses this:

> "...I guess when people have trouble identifying themselves with a society and community, they start creating their own society and community [']when I am in the Netherlands people look at me as a Moroccan guy, and when I am in Morocco people look at me as the Dutch guy. We like to say that the only place where I am really at home is in the place in between." (Soufiane, 23, Amsterdam)

The fact that peers offer a space to 'fit in' and to be accepted is important to note. It means that close friends offer a space where they can develop a positive self-image, and where they feel valued by their friends for who they are as Moroccan, Muslim, Dutch, and, most importantly, as a person.

5.4.2 Peers as Providers of Tangible Resources

Some of the participants' close peers also become key providers of tangible resources. In

---

[25]It is important to notice the tentative language I have used here. While the supportive social mechanisms presented in section 5.4.1 that are shared by the youths and their peers to deal with everyday emotionally and psychologically challenging movements

some cases, peers become important sources of (1) knowledge, skills and information as well as (2) social capital.[25]

Peers as Providers of Skills, Knowledge and Information

Some youths report explicitly or through examples how their peers become important providers of skills and knowledge. First of all, most of the youths and their peers share their cognitive skills in order to face and make sense of their everyday issues. Aya's words want to represent this pattern:

> "Sometimes her advice is my leading path. Because I know she wants the best for me and sometimes you don't see what's good for you and knowing that she sees what's good for me and she never disappointed me...I mostly follow her advice" (Aya, 21, Rotterdam)

Furthermore, some participants report having peers who provide them with their 'practical' skills. For example, Soufiane reports how some of his friends help him out with making videos and producing music:

> "They support me. When I used to have a video shoot or something, my good friends would always show up and make sure everybody works along, think with me..." (Soufiane, 23, Amsterdam)

Other participants indicate how their close friends help them with school. In Aya's words, 'they get me further in school, if I don't get something they explain it to me'. At the same time, some of the participants report helping their peers with the skills that they have.

Lastly, they might also share with each other information that one might not have access to. Omar's story addresses this:

> 'When we grew up he was in the primary school and he told me he was going to this school[...] I just wanted to go to the same school, because he was the only I knew with the same level. [...] I heard

from him he was going to that school [VWO], and I said ok...I also going to sign up for that school." (Omar, 21, Rotterdam)

Peers as Networks

Some youths indicate how their friends were fundamental to providing them with access to institutional agents. For instance, Anas recounted how their friends helped him access his teacher:

> "The toughest moments...when I was about to get kicked off from my VMBO T education because my grades were lower. But they [my friends] really get out me, they really helped me to motivate myself and get my grades back together, quizzing me and also other stuff...." (Anas, 19, Amsterdam)

Similarly, Soufiane tells how his friend put him in contact with a key figure ' who Soufiane addresses as his 'mentor' - in the music industry who is currently helping him in producing his own music:

> "I think I wanna enter the music industry but I have to find an entrance first. That's why I like John that much, because he teaches me everything about it, he puts me in contact with the right people. [...] [John] is my mentor [...] Everyone that's famous right now has been designed by him [...] I met him through a mutual friend who used to work here." (Soufiane, 23, Amsterdam)

In many occasions, peers and close friends were helpful in connecting the participants to institutional agents. In some occasions, the participants' peers used their own networks in the attempt to aid the participants' causes. This was the case of Fatima, whose friends helped her to raise money for the charity project she initiated:

> "Once with my friends...what they did for me. I wanted to raise money for a woman, she was living in Syria and she was a single mom. [...] Then we asked

were found in all of the participants' narratives, here the expression 'in some cases' indicate that not all of the participants have explicitly stated or given examples of how their close friends provide them with tangible resources in terms of knowledge, skills, information and networks.

people. In five days we raised one thousand euros! Oh my God. I am so happy for my friends [...] That's something I will never forget!" (Fatima, 23, Amsterdam)

Furthermore, there are some cases in which the participants and their peers joined their individual networks to create opportunities for their group of friends. For example, Fatima started a non-profit organization aiming at collecting resources for refugee centres in world conflict areas. Another example is Zakaria. Zakaria and his friends joined their resources in order to participate in the party industry. Similarly, Ilyas is starting a business with a friend of his with the help of Ilyas's uncle. This latter example shows how the participants themselves might help each other out by sharing their networks and the resources of their network with their peers.

5.4.3 The Value of Peers

The majority of the participants consider their friends to play an important role in their lives, although not to the same extent. For example, Zakaria openly reports that, while he thinks his friends are important, it is not the case that he 'could not live without them'. This does not seem to be the case for other participants. Aya's words want to indicate how some participants attach great value to their friends:

> "I feel a lot of people have not been there for me. All of my family members, they are there for me, but sometimes they are not [...] I always have to prove them that I do good, I always try to do my best just to get a little, a little happiness just from them [...] but I don't feel I always to be supported by my family [...] And Zineb! She is really my cheerleader! [laugh] yeah, she...is actually always there for me." (Aya, 21, Rotterdam)

The fact that their friends are there when they perceive that no one else is seems to be the reason why they consider their friends to play a vital role in their lives. This might explain why Zakaria openly says that it is not the case that he 'couldn't live without his friends'. It could be that Zakaria

perceives other forms of support. This, however, does not mean that the participants who perceive to be supported by adults ' whether they are parents, role models, teachers, mentors or sport trainers ' do not attach value to their peers. The model of happiness (Figure 14) that Fatima drew when she was asked to 'put her life on paper' seems to show precisely that, even when the participants perceive to be surrounded by adults who support them, friends become important to 'achieve happiness'. This is how Fatima describes her model (Figure 14) and the role of friends:

> "I think family because that's the base, that's where you come from. Friends, you need them outside when you are young [...] then you go to school. but you always have that to fall back on your friends and family". (Fatima 24 Amsterdam)

As Fatima explains, 'you need a strong base' and friends seem to become an important element of such a base.

Discussion

This chapter discusses the findings in light of the research context and the theories employed to frame the study and it does so in three separate sections. The first section discusses the findings presented in section 5.2, which described who the peers of the Moroccan-Dutch youths that participated in the research are and the processes underlying peer group formation as described by the participants. The second section discusses the findings outlined in section 5.3 concerning the role that peers and peer culture play in the process of Moroccan-Dutch youths' identity formation. The third and last section presents the discussion of the findings reported in section 5.4, which identified the types of support that the participants are provided with by their peers. In each section, the limitations of the study are also presented. The last section of the chapter addresses the study's implications for policy and research.

6.1 Peer Group Formation as a Search for Understanding: Beyond the Myth of Self-Segregation

With regards to who the Moroccan-Dutch youths' peers are, the findings reveal that the demographics of their peers depend on the characteristics of the contexts of interaction and more specifically on the participants' neighbourhoods and schools. Since almost all the youths live and were raised in highly diverse neighbourhoods featuring a low share of native Dutch youths - which reflects the residential patterns of Moroccan and Dutch families (van Prag & Schoorl, 2008; Crul & Heering, 2008; Musterd & Fullaondo, 2008; Crul et al., 2013) - the participants seldom come into contact with native Dutch youths. For this reason, secondary education becomes the domain where the youths might come in contact with native Dutch peers, as previous research has already found (Heat et al., 2008; OECD, 2010). However, only the youths attending higher levels of the Dutch high school system report having native Dutch schoolmates, which reflects the segregation of the Dutch educational system (Crul & Doomernik, 2003; Tesser & Dronkers, 2007; Stevens et al., 2011; Turcatti, 2018).

It follows that neighbourhood and school segregation do not facilitate the formation of peer groups that include native Dutch youths, as previous research showed (OECD, 2010), since, as the findings reveal, exposure to contact and geographical proximity are important factors for facilitating the formation of peer groups. Rather, school and neighbourhood segregation facilitate the formation of highly diverse peer groups, insofar as segregation in the Netherlands is characterized by the spatial separation of middle- and upper class native Dutch families, from working-class groups of a diverse background (Crul & Heering, 2008; Muster & Fullaondo, 2008).

The study also reveals that peer groups are formed on the basis of multiple social identities, as indicated by previous research concerning Mexican youths in the US (Haayen, 2016). This is because the youths assume that by sharing specific social identities (e.g. being Moroccans, being second-generation migrants), their peers will also share similar ideas, values, cultural references, and struggles. This suggests that for the participants ethnicity is not the primary and sole classificatory system through which they frame their peer relationships. Rather, by bonding with similar others, the participants seem to seek mutual understanding in their peers. Although these findings are not generalizable, they contribute to deconstructing the myth of Moroccan-Dutch youths' self-segregation that is hegemonic in contemporary Dutch discourse (Turcatti, 2017a).

What is worthwhile to reflect upon is why 'being Dutch' is not considered by the youths a social identity on the basis of which peer groups can form. The study suggests that it is not the case that peer groups are not formed around Dutch identity because they do not identify as Dutch. As a matter of fact, contrary to the branch of research that found an overall tendency to reject Dutch culture among Moroccan-Dutch youths (Maliepaard et al., 2010; Maliepaard, 2012; Maliepaard et al., 2012), the participants actively claim their right to be Dutch. When they do not claim their 'Dutchness', they specify that it is because they feel unwanted, which corroborates the effects that inequality and discrimination have on the youths' Dutch identification (Verkuyten et al., 2012; De Vroome et al., 2014). Rather, the findings indicate that peer groups are not built around Dutch identity because the youths construct being Dutch as being native, socio-economically privileged, Christian, and having false assumptions with regards to those holding a different positionality in terms of background. Interestingly, this construction has much in common with the hegemonic understanding of Dutch identity (Ghorashi, 2016). Hence, feeling excluded from the common-sense construction of Dutch identity seem to explain why the youths do not form peer groups on the basis of being Dutch.

Lastly, in line with the US scholarship on migrant youths' peer group formation (Gibson et al., 2004; Gandara et al., 2004; Stanton-Salazar & Spina, 2005), the present research reveals that peer group formation depends on the outcomes of peer relationship. Hence, the importance of the positive outcome of peer relationships for peer group formation leaves open a window for also bonding with peers with whom the youths might not share any of the social identities identified as salient for peer group formation (e.g. economically privileged native Dutch peers who did not grow in highly diverse neighbourhoods). Furthermore, the fact that being non-judgemental is salient suggests that the youths are to be found at the inter-

sections of different social groups expecting loyalty from them. Indeed, not only they are sometimes accused of not being Dutch 'enough', but they can be also accused of being 'too Dutch' by their Moroccan-Dutch peers. This suggests that also Moroccan-Dutch youths might be liable to be accused of 'acting white' (Fordham & Ogbu, 1986), as Nibbs (2016) found with regards to peer group

formation among Hmong youths in the US, and that the youths view their friendships as a shelter from such accusations.

Figure 9 presents the model that I have built on the basis of the interpretation of the findings in light of the theories employed to frame the study explaining peer group formation.
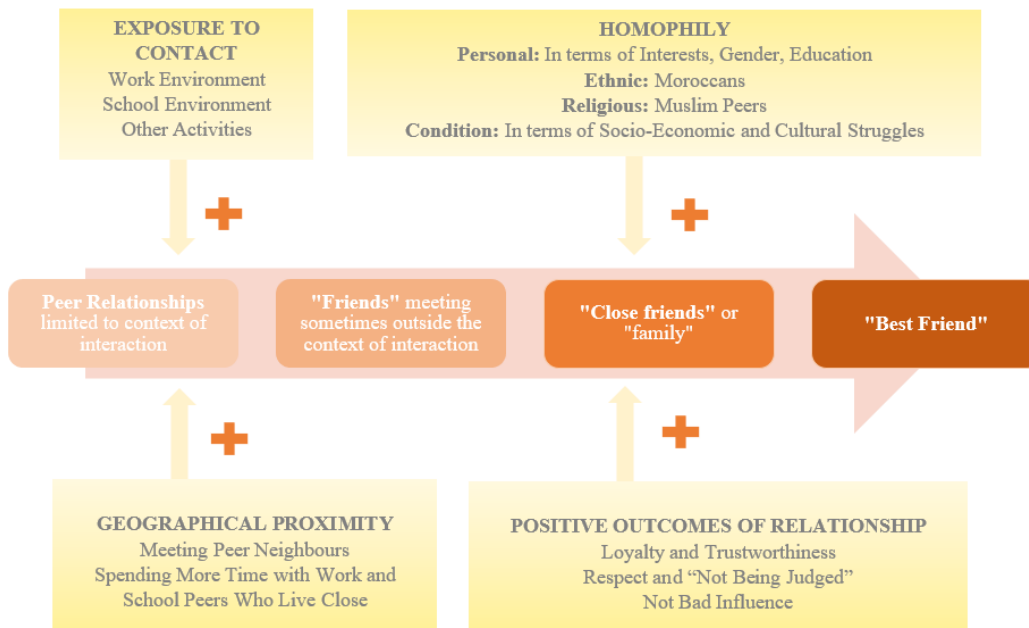
Model Explaining Peer Group Formation among Moroccan-Dutch Youths: exposure to contact, geographical proximity, homophily in terms of different social identities ' but not Dutch identity - and outcomes of peer relationships are the identified factors strengthening peer relationships.

0.1 6.2 More than a Space for Identity-Making: Peer Culture as a Space of Resistance

Concerning the role of peers in the youths' identity-making, the findings are aligned with previous studies showing that peer groups are salient spaces where youths from migrant families construct their identities (Haayen, 2016; Vigil, 2004; Slootman, 2014). While the youths' mundane activities offer spaces to shape their goals and work out decisions about their future in light of their socio-economic position and opportunities, they also become spaces to construct their own understanding of Islam, Dutch, and Moroccan culture. Hence, as Slootman (2014) found concerning Moroccan-Dutch youths' university peers (Slootman, 2014), peer relationships offer a space for acculturation.

Yet, the findings suggest that the youths' accul-

turation ' what Bhabha's (1994) labelled as the 'hybridization' process - is influenced by the youths' individual and collective experiences of inequality and discrimination, which are more strongly suffered by the males. In some cases, the process of making sense of such experiences within the youths' peer groups leads to the unstable, always-shifting construction of a social in-group of Muslims, Moroccans or more generally second-generation youths holding values of hospitality and equality and an out-group of native Dutch where equality and hospitality have no place. Hence, this research seems to suggest that, contrary to Azghari's et al. (2017) results and in line with Slootman's findings (2014), it is not necessarily the case that lack of contact with native-Dutch peers and regular co-ethnic contact prevent the youths from integrating. Rather, it seems that it is discrimination that makes it harder for the youths to endorse integra-

tion as an acculturation strategy, as already theorized by Berry (2005) and revealed by research on Latin American peer groups in the US (Vigil, 2004).

Figure 10 presents the model that I have built on the basis of the interpretation of the findings in light of the theories employed to frame the study explaining how peer groups provide Moroccan-Dutch youths with a space to do identity work.
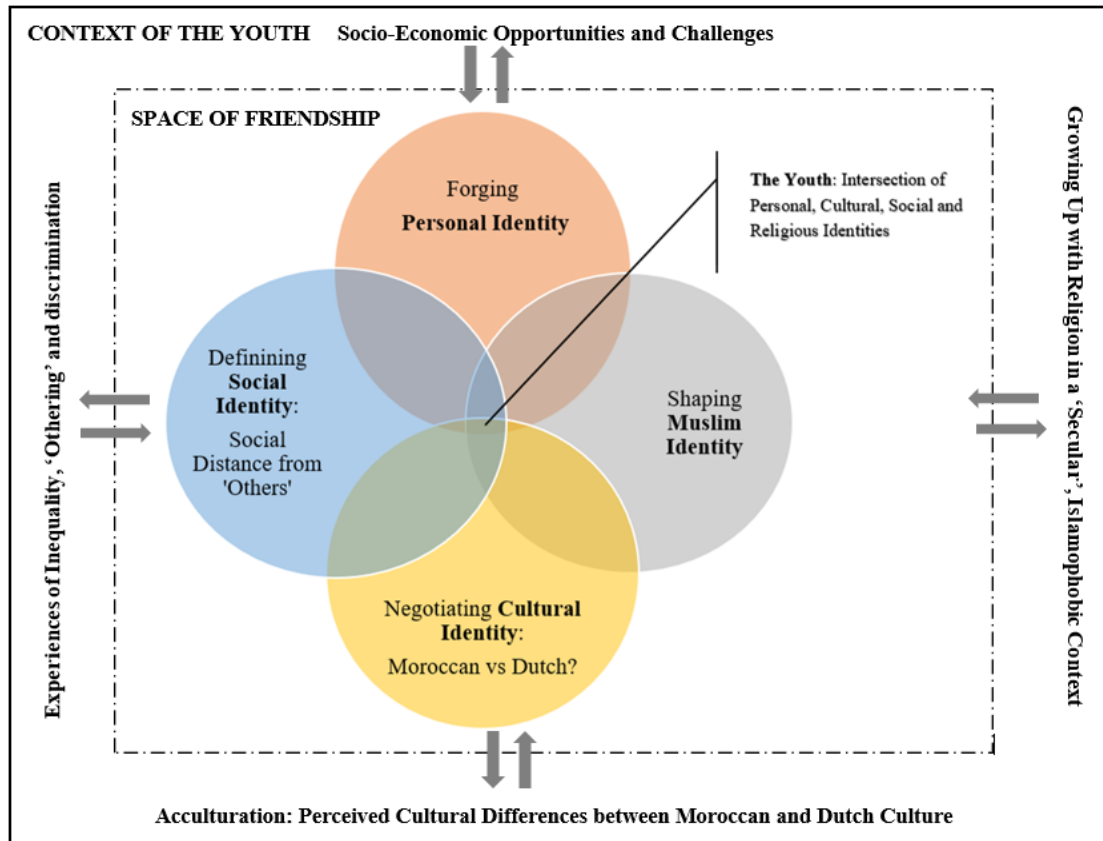
Model Explaining How Peer Groups Become Spaces for Moroccan-Dutch Youths' Identity-Formation and highlighting how the identity work taking place within the space of friendship needs to be contextualized in light of the youths' positionality in socio-economic, cultural, and religious terms.

While peers offer a space for Moroccan-Dutch youths' identity construction, the research also shows that the youths and their friends come to share a set of meanings and norms that regulate not only peer interactions but also everyday practices. The youths and their peers collectively value success in economic terms as they want to achieve a better socio-economic position than their parents, which has been identified as fundamental for the upward socioeconomic mobility of youths from migrant families (Heat et al., 2008). The youths and their peers also highly value reciprocal help and seem to endorse what Stanton-Salazar and Spina (2000) refers to as 'communalistic orientation' (p. 246) which scholars have found to constitute one

of the characteristics of resilient urban minority youths and to be at the basis of such youths' socio-economic mobility (Stanton-Salazar & Spina, 2000; Stanton-Salazar & Spina, 2005; Stanton-Salazar, 2011). Furthermore, the research discovered respect, acknowledgement and openness to diversity and differences to be at the basis of their interpersonal interactions and ways of thinking. These findings suggest that the peer culture that the Moroccan-Dutch youths that participated in the research have established with their peers might promote orientations highly beneficial in the long run. Hence, despite the fact that the research findings cannot be generalized to the whole population of Moroccan-Dutch youths, the research fundamen-

tally opposes the hegemonic constructions of the youths' peers and peer networks as pathological.

However, it is important to notice that the research is limited in the sense that little participant observation was conducted and as such it had little opportunity to investigate how the values and norms that the youths report to share with their peers are embedded in their everyday practices and regulate the interaction within peer groups. In order to do this, I would have needed to master the Dutch language and additional time to conduct the research. Further participant observation would have shed light on the ways the control mechanisms that the youths report to share with their peers are embedded in their practices, the extent to which they are respected, and the instances in which the youths fail to comply with them and with what consequences. Furthermore, the validity of the research would have benefitted from interviewing systematically the members of Moroccan-Dutch youths' peer groups. Indeed, only in some occasions did I interview Moroccan-Dutch youths' belonging to the same peer group. In addition, I did not interview any of the Moroccan-Dutch youths' friends who have a background other than Moroccan, although I sometimes had informal conversations with them.

Despite these limitations, the findings suggest that such peer culture might be regarded as Moroccan-Dutch youths' space of resistance. Indeed, the fact that the youths aim for a higher social status seems to reflect their drive to claim equal status in spite of the diminishing societal characterizations that attempt to 'put them in their place' and might represent for them - in Steele and Aronson's words (1995) - a 'stereotype threat'. By valuing a 'communalistic orientation' (Stanton-Salazar & Spina, 2000, p. 246), the youths seem to resist individualistic explanations of success that can be deployed to justify the actual and potential 'failures' of Moroccan-Dutch youths. In addition, the fact that they embrace diversity seems to indicate how they resist and reject altogether assimilation as a way to position themselves in Dutch society and as a norm to regulate their everyday life relationships. This leads to an openness to *difference*, which the research discovered to be facilitated by *positive* contact with those who are perceived as 'different' - whether they are native Dutch, or of a background other than Moroccan - and not merely by contact in quantitative terms, as

Azghari's et al. (2017) results suggest.

What is worthwhile to reflect upon is the fact that some of the youths explicitly address that it is specifically the peer culture they share with their peers that makes them and their peers ' whether of Moroccan, Dutch, or Surinamese background ' unique and innovative in Dutch society. Hence, these findings seem to suggest that not only peer culture becomes a space for resistance, but also a space to define how they fit in a society characterized by ethnic hierarchies. However, as the findings indicate, some of the youths, despite sharing in the outlined peer culture, seem to still be struggling with proving wrong the hegemonic and diminishing representations, instead of considering them misconceived in the first place, which is what was found to protect stigmatized individuals from the effects of the stereotype threat (Schmader, 2010; Kinias & Sim, 2011; Spencer et al., 2016; Vermeulen et al., 2016). This seems to take away much of the empowerment that such critical awareness brings. Whether the youths are able to identify such hegemonic representations as misconceived in the first place seems to depend on the level of their own self-confidence which appears to be tied to the presence of adults and opportunities that help them to develop a sense of self-worth. Ultimately, these findings not only point to the importance of structural opportunities to develop such self-confidence, but also the saliency of adults in helping the youths in the process.

6.3 Peer Support and Socio-Economic Mobility: A Necessary *and* Sufficient 'Back-Up"

With regards to the role of peers as providers of support, the research reveals that the Moroccan-Dutch youths that participated in the research and their peers contribute to each other's psychological and emotional wellbeing, as research also found concerning the role of peers in the lives of migrant youths in the US (Stanton-Salazar & Spina, 2000; Suarez-Orozco & Todorova, 2003; Stanton-Salazar & Spina, 2005). They do so by providing each other with two important types of support outlined in House's (1981) social support typology, namely emotional support and support in re-appraising stressful events. Indeed, the youths and their peers establish always-accessible safe spaces where issues can be talked about freely

or where it is possible to distract themselves. At times they support each other's ideas, while also challenging each other's evaluation of specific situations (Figure 18). This kind of support aids the youths to cope with stress, address stress and even re-evaluate the extent to which a situation is perceived as stressful. For these reasons, also Moroccan-Dutch youths' peers might be regarded as protective factors, as scholars already discovered with regards the peers of working-class youths from migrant families in the US (Stanton-Salazar & Spina, 2005; Stanton-Salazar, 2011). These findings seem to suggest that the youths are able to provide each other with these two kinds of support independently of their class, educational status and background. This is an important remark to be made given the pathologizing of the peer networks of working-class Moroccan-Dutch youths (Jurgens, 2007).

Furthermore, the findings reveal that not only do peers contribute to each other's psychological and emotional well-being, but that in some occasions peers become vehicles of social capital and crucial knowledge (Figure 18). Hence, Moroccan-Dutch youths' peers can also provide the other two types of support that characterize House's (1981) social support typology ' instrumental and informational support - which in sociological theory have been indicated as fundamental for socio-economic mobility (Bourdieu, 1977; Stanton-Salazar, 2004). Indeed, some of the participants report to relying on their peers' skills in a given expertise and capacity for critical thinking. In some occasions, the youths also report how their peers share crucial information with them about labour and educational opportunities in addition to social contacts. It is important to notice that these peers were not necessarily native Dutch or middle class. While this deconstructs the assumption that Moroccan-Dutch youths can only benefit from contact with middle class and/or native Dutch peers (Azghari et al., 2017), they also support Stanton-Salazar's (2004) observation regarding how working-class youths from migrant families can also be embedded in social sites that provide access to resources and institutional agents.

However, it is important to notice that while some of the participants reported multiple examples of how their peers share social capital, skills, and information with them, some of the youths did not provide many examples of this. Although ab-

sence of evidence is not evidence of absence, it might be the case that not all of the participants' close peers have skills, knowledge and social resources to share, which might be indeed related to the position of their peers in terms of class and education, as scholars might suggest (Stanton-Salazar, 2004; Stanton-Salazar, 2011). Furthermore, the fact that some youths place greater value in their peers than others seems to be due the fact that such youths perceive peer support as one among the few kinds of support they receive by other social actors on a daily basis. It is interesting to notice that it was not necessarily the highly educated who place less value in their peers, but those who reported to have mentors, role models and loving adults within their families and the various communities in which they are embedded. If the hypothesis that the youths placing greater value in their peers are those who have limited access to support from other social actors is valid, it makes us wonder whether the support of their peers is not just necessary but sufficient to support the youths to withstand challenging moments and the socio-economic mobility of the youths who do not access social capital and information through their peers or other channels.

0.2   6.4 Implications & New Directions: Decolonizing Policy Paradigms and Research Practices

The analysis of the youths' narratives addressing the role that peers have in their lives have implications for policy and research, which are presented in the following subsections.

6.4.1 From Assimilationist, Colour-Blind Policies to Bottom-Up, Intercultural Policies

The research indicates that there is a need to decolonize policy paradigms in order to enhance the opportunity structure of Moroccan-Dutch youths, which would benefit not only the youths' socio-economic mobility, but also their acculturation process and self-confidence building, as the research indicates. Policies should not be formulated on the basis of cultural explanations of failure currently underlying assimilationist frameworks, since such explanations problematize the youths' heritage, culture, and religion and lead to formulation of policies that aim to change the youths' cul-

tural and religious practices (Montero-Sieburth & Perez, 2014). Instead, policies should strive to enhance equity by dismantling structural inequality. It is therefore necessary for policies to be formulated on the basis of evidence. As such, policies need to be based on research addressing what Moroccan-Dutch youths, their communities and the adults indicate to be detrimental for Moroccan-Dutch youths' functioning, rather than on what it is thought that they need.

Furthermore, the research shows that there is a need for policies promoting intercultural dialogue and respect, which at the moment seems to be Moroccan-Dutch youths' burden rather than a shared responsibility. This requires a fundamental shift away from the current assimilationist framework (Schinkel, 2013; Wekker, 2016) and the adoption of an intercultural paradigm. Such a framework should not essentialize cultural differences, but promote mutual understanding and replace tolerance with respect, in addition to encouraging a critical awareness of one's own positionality and cultural hybridity. Furthermore, such an intercultural framework could decrease the emphasis within Dutch society on cultural differences. As the participants of the research taught us, acknowledging the commonalities that exist between individuals and groups differently positioned in terms of culture, socio-economic background or religion enables mutual understanding beyond social divisions. If accompanied by the desegregation of native Dutch spaces, such intercultural dialogue not only would facilitate the youths' process of acculturation in social, cultural, and psychological terms, but it would also ease inter-ethnic tensions.

The research also suggests the need for implementing programs at the community level encouraging the involvement of adults in the lives of Moroccan-Dutch youths. Training institutional agents in understanding the specific challenges that Moroccan-Dutch youths face in Dutch society might inform them on how to bridge the youths to socio-economic and cultural resources. Furthermore, initiatives are needed which are inclusive of youths, parents, and adults with whom the youths are in contact on a regular basis (e.g. teachers) aiming to strength the social networks and support of the youths and enhance the intergenerational understanding that some youths seem to lack. There is also the need for programs encouraging peer support and knowledge circulation among peer networks. More radically, programs aimed at building the self-confidence of the youths and featuring Moroccan-Dutch role models of social mobility are needed. Such initiatives could help the youths to develop the critical thinking and sense of self-worth that the research discovered to be fundamental to cope with and flourish despite the hegemonic characterizations embedded in everyday life and diminishing the youths' potential and worth.

6.4.2 From Researching Deviance to Exploring Youths' Strengths and Meaning-Making

The study also suggests that there is a need of decolonizing research practices by focusing on the youths' strengths. Indeed, the study shows how such a focus enables to counteract the hegemonic representations of Moroccan-Dutch youths as lacking, deviant, and failing, while doing justice to the youths' successes and contributions to Dutch society. Furthermore, this study demonstrates that focusing on the youths' strengths enables to identify structures that limit their functioning and that can be changed, insofar as it enables to locate failure and deviance in broader structures of inequalities. As such, it problematizes inequality rather the youths' culture, family, or networks, while allowing to identify beneficial socio-cultural mechanisms that can be strengthened through intervention.

Furthermore, this study shows how qualitative research conducted with an interpretive and critical approach is needed to complement the vast body of longitudinal, quantitative studies locating the socio-economic position of Moroccan-Dutch youths in Dutch society. Indeed, the study demonstrates how this methodological approach not only allows us to address the often-dismissed perspectives of the youths, but also to understand their own perspectives on what needs to be changed, while decolonizing the theoretical explanations through the youths' own critical thinking. More fundamentally, the decolonizing of theoretical explanations would largely benefit if research were to be conducted by Moroccan-Dutch scholars themselves. Indeed, as Darder (2018) writes when addressing the need to decolonize interpretive research, 'the sensibilities of subaltern voices can offer epistemological breakthroughs necessary to forging transformative political praxis linked to bringing together decolonizing theories and practices' (p. 99).

This approach is needed in order to expand the knowledge on the sociocultural mechanisms and meaning-making that the youths deploy in order to navigate Dutch society. While this research addressed the role that peers play in Moroccan-Dutch youths' navigation in Dutch society and how Moroccan-Dutch youths and their peers make sense of themselves and their own position in society, much remains unknown of belief systems developed by the youths. Critical and interpretive research might want to consider taking as a unit of analysis Moroccan-Dutch youths' families and investigate the intergenerational dynamics in terms of parental participation in the youths' lives and the transmission of parental beliefs to the children and whether these enable to the youths to develop a sense of who they are and who they want to be. Also necessary are community-based studies that address the intersection of the youths' families, communities and schools in order to better understand the youths' supportive system and the extent to which these domains cooperate to favour the youths' flourishment.

7 Conclusion

This critical ethnography investigated the meaning-making that thirty Moroccan-Dutch youths share with their peers and deploy to navigate Dutch society. The research was informed by an interdisciplinary theoretical framework and was based on 110 hours of fieldwork conducted between January 2017 and May 2018, during which multiple forms of qualitative data were collected. The research wanted to unveil the strengths of the youths' peer culture and networks in the attempt to move beyond the pathologizing of the youths and discover positive peer mechanisms that can be supported. As such, the research wanted to (1) identify who the peers of Moroccan-Dutch youths are and what mechanisms the youths identify at the basis of peer group formation; (2) understand the role of peers and peer culture in the process of identity-formation of Moroccan-Dutch youths; (3) identify the kinds of support that peers provide to Moroccan-Dutch youths to face everyday challenges and to navigate Dutch society.

The research revealed that the formation of highly diverse peer groups that not always include native Dutch youths is shaped by segregation and by a hegemonic and exclusive construction of Dutch identity and is underlined by the youths' search for understanding and safe space from everyday discrimination. Such peer groups become spaces where Moroccan-Dutch youths forge their personal, social and cultural identities, while constructing a peer culture that promotes success in socio-economic terms, communalistic orientations, embracement of diversity, and openness to differences, and in which the youths who do not have access to other supportive networks rely for their psychological and emotional wellbeing, critical information, and social capital. However, peers are not always the best access to opportunity and such peer culture is difficult to maintain in light of discriminatory practices and limited opportunity structure. The youths with a strong sense of self-worth, which is tied to the presence of adults and opportunities, and who recognize the innovativeness of the peer culture that they share with their peers are the most successful to maintain such orientations and to appropriate their space in Dutch society, despite persisting discrimination and lack of recognition.

While showing how peers should not be devalued as resources and spaces for development, this study challenges the hegemonic construction of Moroccan-Dutch youths' as pathological, unintegrated and unwilling to do so. Although further research is needed to understand the belief systems developed by the youths in relation to intergenerational dynamics and community embeddedness, this research showed that investigating the youths' strengths with an interpretive and critical approach enables to move beyond the pathologizing of the youths, to place failure in broader structures of oppression, while decolonizing theoretical explanations through the youths' critical thinking. More fundamentally, the study indicated the urgency for evidence-based, intercultural policies dismantling structural inequality, encouraging supportive networks, providing role models for social mobility, while promoting intercultural respect.

If this is not done, we might ask ourselves what the price to pay will be. Moroccan-Dutch youths will have to keep doing 'their own little thing' under harsh circumstances, as they will still not be 'accepted in the first place'. Some of them will be able to appropriate their space in Dutch society, but others will not. While this represents an infringement of the right for self-development and affirmation of

one's full potential, it represents a great loss for Dutch society, which has a great deal to learn from the peer culture that Moroccan-Dutch youths share with their peers which, given the context in which it develops, might be regarded as futuristic. Ultimately, the question is whether we want to embrace or not the future: *'c'mon, dude, this is the twenty-first century'.*

References

Andriessen, I., Nievers, E., Dagevos, J., & Faulk, L. (2012). Ethnic Discrimination in the Dutch Labor Market: Its Relationship with Job Characteristics and Multiple Group Membership. Work and Occupations, 39(3), 237'269. DOI: 10.1177/0730888412444783

Azghari, Y., van de Vijver, F. J. R., & Hooghiemstra, E. (2017). Identity as a Key Factor in the Acculturation of Youth Moroccan-Dutch Adults. Journal of Psychology in Africa, 27(2), 132-140. DOI: 10.1080/14330237.2017.1301706

Bank, B. J. (1997). Peer Cultures and Their Challenges for Teaching. In B. J. Biddle, T. L. Good, I. Goodson, (Eds.), International Handbook of teachers and Teaching. (pp. 879-937). New York, NY: Springer.

Barker, C. (2004). The SAGE Dictionary of Cultural Studies. London, UK: SAGE Publications.

Berry, J. W. (2005). Acculturation: Living Successfully in Two Cultures. International Journal of Intercultural Relations, 29(6), 697'712. DOI: 10.1016/j.ijintrel.2005.07.013

Bhabha, H. K. (1994). The Location of Culture. London, UK: Routledge.

Bourdieu, P. (1977). Cultural Reproduction and Social Reproduction. In J. Karabel & A. H. Hasley (Eds.), Power and Ideology in Education. (pp. 487-511). New York, NY: Oxford University Press.

Bourdieu, P. (1986). The Forms of Capital. In J. G. Richardson (Ed.), Handbook of the Theory and Research for the Sociology of Education. (pp. 241-258). New York, NY: Greenwood Press.

Brettell, C. B., & Nibbs, F. G. (2009). Lived Hybridity: Second-Generation Identity Construction through College Festival. Identities, 16(6), 678-699. DOI: 10.1080/10702890903307142

Buijs, F. J. (2009). Muslims in the Netherlands: Social and Political Developments after 9/11. Journal of Ethnic and Migration Studies, 35(3), 421-438. DOI: 10.1080/13691830802704590 Campbell, E., & Lassiter, L. E. (2015). Doing Ethnography Today: Theories, Methods, Exercises. Oxford, UK: Wiley Blackwell.

Carspecken, P. F. (1996). Critical Ethnography in Educational Research. New York, NY: Routledge.

CBS (2012). Annual Report on Integration 2012: Summary. CBS Retrieved from: http://www.cbs.nl/NR/rdonlyres/CF430FF1-4623-4D60-B34B-CBE33749006A/0/2012b72pub.pd

CBS (2016). Annual Report on Integration 2016: Summary. CBS Retrieved from: https://www.cbs.nl/en-gb/publication/2016/47/annual-report-on-integration-2016

Celeste, L., Meeussen, L., Verschueren, K., & Phalet, K. (2016). Minority Acculturation and Peer Rejection: Costs of Acculturation Misfit with Peer-Group Norms. British Journal of Social Psychology, 55, 544-563. DOI:10.1111/bjso.12147

City of Rotterdam Regional Steering Committee (2009). The City of Rotterdam, The Netherlands: Self-Evaluation Report. In OECD, Reviews of Higher Education in Regional and City Development. Retrieved from: http://www.oecd.org/edu/imhe/regionaldevelopmen

Cohen, S., & McKay, G. (1984). Social Support, Stress and the Buffering Hypothesis: A Theoretical Analysis. In A. Baum, S. E. Taylor, & J. E. Singer (Eds.), Handbook of Psychology and Health. (pp. 253-267). Erlbaum, NJ: Hillsdale.

Cohen, S., & Syme, S. L. (1985). Issue in the Study and Application of Social Support. In S. Cohen and S.L. Syme (Eds.), Social Support and Health (pp. 3'22). Orlando, FL: Academic Press Inc.

Corsaro, W. A. (1988). Routines in the Peer Culture of American and Italian Nursery School Children. Sociology of Education, 61(1), 1-14. Retrieved from: URL: http://www.jstor.org/stable/2112305

Corsaro, A. W., & Eder, D. (1990). Children's Peer Cultures. Annual Review of Sociology, 16, 197-202. DOI: 0360-0572/9010815-

019702.00

Crul, M., & Doomernik, J. (2003). The Turkish and Moroccan Second Generation in the Netherlands: Divergent Trends between and Polarization within the Two Groups. International Migration Review, 37(4), 1039'1064. DOI: 10.1111/j.1747-7379.2003.tb00169.x

Crul, M., & Heering, L. (2008). The Position of the Turkish and Moroccan Second Generation in Amsterdam and Rotterdam: The TIES study in the Netherlands. IMISCOE Research: Amsterdam University Press.

Crul, M., Schneider, J., & Leile, F. (2013). Super-Diversity: A New Perspective on Integration. Amsterdam, the Netherlands: VU University Press.

Darder, A. (2018). Decolonizing Interpretive Research: Subaltern Sensibilities and The Politics of Voice. Qualitative Research Journal, 18(2), 94-104. DOI 10.1108/QRJ-D-17-00056 de Jong, J. D. A. (2012). Typically Moroccan' A Group Dynamic Explanation of Nuisance and Criminal Behavior. In F. A. Esbensen & C. L. Maxson (Eds.), Youth Gangs in International Perspective. Results from the Eurogang Program of Research (pp. 225-236). New York: Springer

De Vroome, T., Verkuyten, M., & Martinovic, B. (2014). Host National Identification of Immigrants in the Netherlands. International Migration Review, 48(1), 1'27. DOI: 10.1111/imre.12063

Demanet, J., & Van Houtte, M. (2012). School Belonging and School Misconduct: The Differing Role of Teacher and Peer Attachment. Journal of Youth Adolescence, 41, 499'514. DOI: 10.1007/s10964-011-9674-2

Dekovic, M., Wissink, I. B., & Marie, A. M. (2004). The Role of Family and Peer Relations in Adolescent Antisocial Behaviour: Comparison of Four Ethnic Groups. Journal of Adolescence, 27, 497'514. DOI: 10.1016/j.adolescence.2004.06.010

Di Stasio, V., & Van De Werfhorst, H. G. (2016). Why does Education Matter to Employers in Different Institutional Contexts' A Vignette Study in England and the Netherlands. Social Forces, 95(1), 77'106. DOI: 10.1093/sf/sow027

Eijberts, M., & Roggeband, C. (2016). Stuck with the Stigma' How Muslim Migrant Women in the Netherlands Deal - Individually and Collectively - with Negative Stereotypes. Ethnicities, 16(1), 130'153. DOI: 0.1177/1468796815578560

Engbersen, G., Leerkes, A., Grabowska-Lusinska, I., Snel, E., & Burgers, J. (2013). On the Differential Attachments of Migrants from Central and Eastern Europe: A Typology of Labour Migration. Journal of Ethnic and Migration Studies, 39(6), 959-981. DOI: 10.108/1369183X.2013.765663

Essed, P. (1996). Diversity, Gender, Color & Culture. Amherst, Massachusetts. University of Massachusetts Press.

Essed, P. & Hoving, I. (2014). Innocence, Smug Ignorance, Resentment: An Introduction to Dutch Racism. In P. Essed & I. Hoving (Eds.), Dutch Racism. Amsterdam, The Netherlands: Rodopi.

Fine, M. (1994). Dis-tance and Other Stances: Negotiations of Power Inside Feminist Research. In A. Gitlin (Ed.), Power and Method. New York, NY: Routledge.

Flick, U. (2014). An Introduction to Qualitative Research (5th ed.). New York, NY: SAGE Publications

Fordham, S., & Ogbu, J. (1986). Black Students' School Success: Coping with the 'Burden of 'Acting' White'. Urban Review, 18(3), 176-206. Retrieved from: https://link-springer-com.proxy.uba.uva.nl:2443/content/pdf/10.10072FBF01112192.pdf

Forum (2010). The Position of Muslim in the Netherlands: Facts and Figures. Fact book. Institute for Multicultural Affairs. Retrieved from: http://www.eukn.eu/fileadmin/Lib/files/EUKN/2010/0_linkclick.pdf

Futch, V. A., & Fine, M. (2014). Mapping as a Method: History and Theoretical Commitments. Qualitative Research in Psychology, 11(1), 42-59. DOI: 10.1080/14780887.2012.719070

Gándara, P., O'Hara, S., & Gutierrez, D. (2004). The Changing Shape of Aspirations: Peer Influence on Achievement Behaviour. In M. A. Gibson, P. Gándara & J. P. Koyama (Eds.), School Connections: U.S. Mexican Youth, Peers, and School Achievement. (pp. 39-62). New York, NY: Teachers College Press.

Geertz, C. (1973). The Interpretation of Cultures.

New York, NY: Basic Books.

Ghorashi, H. (2016). Ways to Survive, Battles to Win: Iranian Women Exiles in the Netherlands and the United States. New York, NY: Nova Science Publishers.

Gibson, M. A., Gándara, P., & Koyama, J. P. (2004). The Role of Peers in the Schooling of U.S. Mexican Youth. In M. A. Gibson, P. Gándara & J. P. Koyama (Eds.), School Connections: U.S. Mexican Youth, Peers, and School Achievement. (pp. 1-17). New York, NY: Teachers College Press.

Giroux, H. (1981). Ideology, Culture & The Process of Schooling. Philadelphia, Pennsylvania: Temple University Press.

González, F., Moskowitz, A., & Casto-Gómez, S. (2001). Traditional vs. Critical Cultural Theory. Cultural Critique, 49, 139-154. Retrieved from: www-jstor.org/stable/1354706

Gracia, P., Vázquez-Quesada, L., & Van de Werfhorst, H. G. (2016). Ethnic Penalties' The Role of Human Capital and Social Origins in Labour Market Outcomes of Second-Generation Moroccans and Turks in the Netherlands. Journal of Ethnic and Migration Studies, 42(1), 69'87. DOI: 10.1080/1369183X.2015.1085800

Granovetter, M. S. (1973). The Strength of Weak Ties. The American Journal of Sociology, 78(6), 1360-1380. Retrieved from: http://www.jstor.org/stable/2776392

Grillo, R. D. (2003). Cultural Essentialism and Cultural Anxiety. Anthropological Theory, 3(2), 157-173. Retrieved from: 10.1177/1463499603003002002

Güngör, D., Fleischmann, F., Phalet, K., & Maliepaard, M. (2013). Contextualizing Religious Acculturation: Cross-Cultural Perspectives on Muslim Minorities in Western Europe. European Psychologist, 18(3), 203'214. DOI: 10.1027/1016-9040/a000162

Haayen, L. (2016). 'My Friends Make Me Who I Am': The Social Spaces of Friendship among Second-Generation Youth. In F. Nibbs & C. B. Brettell (Eds.), Identity and the Second Generation: How Children of Immigrants Find Their Space. (pp. 65-86). Nashville, Tennessee: Vanderbilt University Press.

Hall, S. (1991). Old and New Identities. In A. D. King, Culture, Globalization and the World-System. Basingstoke, UK: Macmillan.

Heath, A. F., Rothon, C. & Kilpi, E. (2008). The Second Generation in Western Europe: Education, Unemployment, and Occupational Attainment. Annual Review of Sociology, 34, 211-235. DOI: 10.1146/annurev.soc.34.040507.134728

House, J. S. (1981). Work Stress and Social Support. London, UK: Addison-Wesley Publishing Company Longman.

Huijnk, W. & Andriessen, I. (2016). Integratie in zicht' De integratie van migranten in Nederland op acht terreinen nader bekeken. The Hague, the Netherlands: The Netherlands Institute for Social Research.

Jie, D. (2011). Discourse, Identity, and China's Internal Migration: The Long March to the City / Encounters. Bristol, UK: Multilingual Matters.

Jurgens, F. (2007). Het Marokkanendrama. Amsterdam, The Netherlands: Meulenhoff.

Katsiaficas D., Futch, V. A., Fine, M., & Sirin, S. R. (2011). Everyday Hyphens: Exploring Youth Identities with Methodological and Analytical Pluralism. Qualitative Research in Psychology, 8(2), 120-139. DOI: 10.1080/14780887.2011.572743

Keesing, R. M. (1974). Theories of Culture. Annual Review of Anthropology, 3, 73-97. DOI: 10.1146/annurev.an.03.100174.000445

Kinias, Z., & Sim, J. (2016). Facilitating Women's Success in Business: Interrupting the Process of Stereotype Threat through Affirmation of Personal Values. Journal of Applied Psychology, 101(11), 1585-1597. DOI: 10.1037/apl0000139

Lazarus, R.S., & Folkman, S. (1984). Stress, Appraisal and Coping. New York, NY: Springer.

Lewis-Charp, H. Yu, H. C., & Friedlaender, D. (2004). The Influence of Intergroup Relations on School Engagement: Two Cases. In M. A. Gibson, P. Gandara & J. P. Koyama (Eds.), School Connections: U.S. Mexican Youth, Peers, and School Achievement. (pp. 107-128). New York, NY: Teachers College Press.

Lindo, F. (2008). Interetnische Contacten Tussen cholieren in het voVortgezet Onderwijs. Amsterdam, the Netherlands: Institute of Migration and Ethnic Studies.

Madison, D. S. (2012). Critical Ethnography:

Method, Ethics, and Performance (2nd ed.). London, UK: SAGE Publications.

Maliepaard, M. (2012). Religious Trends and Social Integration: Muslim Minorities in the Netherlands. (Doctoral Dissertation). Retrieved from: http://dspace.library.uu.nl:8080/handle/1874/241972

Maliepaard, M., Gijsberts, M., & Lubbers, M. (2012). Reaching the Limits of Secularization' Turkish- and Moroccan-Dutch Muslims in the Netherlands 1998-2006. Journal for the Scientific Study of Religion, 51(2), 359'367. DOI: 10.1111/j.1468-5906.2012.01647.x

Maliepaard, M., Lubbers, M., & Gijsberts, M. (2010). Generational Differences in Ethnic and Religious Attachment and their Interrelation. A Study among Muslim Minorities in the Netherlands. Ethnic and Racial Studies, 33, 451'472. DOI: 10.1080/01419870903318169

Masten, A. S. (1994). Resilience in Individual Development: Successful Adaptation Despite Risk and Adversity. In M. C. Wang & E. W. Gordon (Eds.), Educational Resilience in Inner-City America: Challenges and Prospects. (pp. 3-25). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mignolo, W. D. (2011). The Global South and World Dis/Order. Journal of Anthropological Research, 67(2), 165-18. Retrieved from: http://www.jstor.org/stable/41303282

Mitrousi, S., Travlos, A., Koukia, E., & Zyga, S. (2013). Special Paper: Theoretical Approaches to Coping. International Journal of Caring Sciences, 6 (2), 131-137. Retrieved from: www.internationaljournalofcaringsciences.org

Montero-Sieburth, M., & Perez, L. C. (2014). Recommendations from a Comparative Analysis of Educational Policies and Research for the Achievement of Latinos in the United States and Latin Americans in Spain: Towards Smarter Solutions. In P. R. Portes, S. Salas, P. Baquedano-Lopez, & P. J. Mellom (Eds.), U.S. Latinos and Education Policy: Research-Based Directions for Change. (pp. 92-114). New York, NY: Routledge.

Musterd, S., & Fullaondo, A. (2008). Ethnic Segregation and the Housing Market in two Cities in the Northern and Southern Europe: The Cases of Amsterdam and Barcelona. ACE: Architecture, City and Environment, 3(8), 93-115. Retrieved from: http://www.raco.cat/index.php/ACE

Nibbs, F. G. (2016). 'Too White and Didn't Belong': The Intra-Ethnic Consequences of Second-Generation Digital Diasporas. In F. G. Nibbs & C. B. Brettell (Eds.), Identity and The Second Generation: How Children of Immigrants Find their Space. (pp. 87-103). Nashville, Tennessee: Vanderbilt University Press.

OECD (2010). Equal Opportunities' The Labour Market Integration of the Children of Migrants. OECD. Retrieved from: https://www.oecd-ilibrary.org/social-issues-migration-health/equal-opportunities_9789264086395-en

Pennix, R. (1979). Naar een Algemeen Etnisch Mindrhedembeleid' Wetenschappelijke Raad voor het Regeringsbeleid report. Retrieved from: https://www.wrr.nl/publicaties/rapporten/1979/05/09/minderheden

Peters, K. (2011). Living Together in Multi-Ethnic Neighborhoods: The Meaning of Public Spaces for Issues of Social Integration. Wageningen, The Netherlands: Wageningen Academic.

Pinkster, F. M., & Fortuijn, J. D. (2009). Watch Out for the Neighbourhood Trap! A Case Study on Parental Perceptions of and Strategies to Counter Risks for Children in a Disadvantaged Neighbourhood. Children's Geographies, 7(3), 323-337. DOI: 10.1080/14733280903024498

Poort, A. (2012). Statistiek Saai' CBS-Cijfers Komen Tot Leven Op De Kaart. NRC.nl. Retrieved from: https://www.nrc.nl/nieuws/2012/02/14/statistiek-saai-cbs-cijfers-komen-tot-leven-op-een-kaart-a1447709

Rezai, S., Crul, M., Severiens, S., & Keskiner, E. (2015). Passing the Torch to a New Generation: Educational Support Types and the Second Generation in the Netherlands. Comparative Migration Studies, 3(12), 1-17. DOI 10.1186/s40878-015-0011-x

Rumbaut, R. G. (2004). Ages, Life Stages, and Generational Cohorts: Decomposing the Immigrant First and Second Generations in the United States. International Mi-

gration Review, 38(3), 1160-1205. DOI: 10.1111/j.1747-7379.2004.tb00232.x

Sandoval, L. (2008). Integration and Development Trajectories: Latin American populations in the Netherlands: Exploratory Study. Working Paper 461. The Hague, Netherlands: Institute of Social Studies.

Schinkel, W. (2013). The Imagination of 'Society' in Measurements of Immigrant Integration. Ethnic and Racial Studies, 36(7), 1142-1161. DOI: 10.1080/01419870.2013.783709

Schmader, T. (2010). Stereotype Threat Deconstructed. Current Directions in Psychological Science, 19(1), 14-18. DOI: : 10.1177/0963721409359292

Schwartz, S. J., Montgomery, M. J. & Briones, E. (2006). The Role of Identity in Acculturation among Immigrant People: Theoretical Propositions, Empirical Questions, and Applied Recommendations. Human Development, 49(1), 1-30. DOI: 10.1159/000090300

Shin, R., Daly, B., & Vera, E. (2007). The Relationships of Peer Norms, Ethnic Identity, and Peer Support to School Engagement in Urban Youth. Professional School Counseling, 10(4), 379-388. Retrieved from: http://www.jstor.org/stable/42732576

Shumaker, S. A. & Brownell, A. (1984). Toward a Theory of Social Support: Closing Conceptual Gaps. Journal of Social Issues, 40(4), 11-36. DOI: 10.1111/j.1540-4560.1984.tb01105.x

Slootman, M. (2014). Reinvention of Ethnic Identification among Second Generation Moroccan and Turkish Dutch Social Climbers. New Diversities, 16(1), 56-70. Retrieved from: http://hdl.handle.net/11245/1.447848

Spencer, S. J., Logel, C., & Davies, P. G. (2016): Stereotype Threat. Annual Review of Psychology, 67, 415-437. DOI: 10.1146/annurev-psych-073115-103235

Stanton-Salazar, R. D. (2011). A Social Capital Framework for the Study of Institutional Agents and Their Role in the Empowerment of Low-Status Students and Youth. Youth & Society, 43 (3), 1066-1109. DOI: 10-1177/0044118X10382877

Stanton-Salazar, R. D. (2004). Social Capital among

Working-Class Minority Students. In M. A. Gibson, P. Gandara & J. P. Koyama (Eds.), School Connections: U.S. Mexican Youth, Peers, and School Achievement. (pp. 18-38). New York, NY: Teachers College Press.

Stanton-Salazar, R. D., & Spina, S. U. (2005). Adolescent Peer Networks as a Context for Social and Emotional Support. Youth & Society, 36(4), 379-417. DOI: 10.1177/0044118X04267814

Stanton-Salazar, R. D., & Spina, S. U. (2000). The Network Orientations of Highly Resilient Urban Minority Youth: A Network-Analytic Account of Minority Socialization and Its Educational Implications. Urban Review, 32(3), 227'261. DOI: 10.1023/A:1005122211864

Steele, C. M., & Aronson, J. (1995). Stereotype Threat and the Intellectual Test Performance od African Americans. Journal of Personality and Social Psychology, 69(5), 797-811. Retrieved from: http://users.nber .org/~sewp/events/2005.01.14/Bios+Links /Good-rec2-Steele_&_Aronson_95.pdf

Stevens, P. a. J., Clycq, N., Timmerman, C., & Van Houtte, M. (2011). Researching Race/Ethnicity and Educational Inequality in the Netherlands: a Critical Review of the Research Literature between 1980 and 2008. British Educational Research Journal, 37(1), 5'43. DOI: 10.1080/01411920903342053

Suàrez-Orozco, C., & Todorova, I. L. G. (2003). New Direction for Youth Development: Understanding the Social Worlds of Immigrant Youth. San Francisco, California: Jossey-Bass

. Tesser, P., & Dronkers, J. (2007). Equal Opportunities or Social Closure in the Netherlands. Proceedings of the British Academy 137, 359'401. DOI: 10.5871/bacad/9780197263860.003.0009

Tracy, S. J. (2013). Qualitative research Methods: Collecting Evidence, Crafting Analysis, Communicating Impact. Oxford, UK: Wiley-Blackwell.

Turcatti, D. (2017a). Acculturation as a Matter of Negotiation: Deconstructing the Dutch Myth of Second-Generation Turks and Moroccans' Self-Exclusion. [Unpublished Manuscript].

Turcatti, D. (2017b). Beyond Recognition: Learn-

ing from the Resilience(s) of Moroccan-Dutch Youth. Paper presented at the Who's in Town of Amsterdam University College 'Recognising the resilience of Moroccan youth in the Netherlands: from their perspectives'. Amsterdam, The Netherlands.

Turcatti, D. (2017c). Different Outcomes, Equal Opportunities' The Performance of Second Generation Moroccans in the Dutch Labour Market. [Unpublished Manuscript].

Turcatti, D. (2017d). From Pathological to Supportive Social Relations and Networks: Understanding Social Support's Roles, Values and Meanings in the Everyday Lives of Young Second Generation Moroccans Living in The Netherlands. [Unpublished Manuscript].

Turcatti, D. (2017e). 'I am not only the Muslim woman': Constructing Hybrid Identities and Articulating Religious Agency through Clothes in a Home Country which is 'mine also, as it is yours'. [Unpublished Manuscript].

Turcatti, D. (2017f). Normalizing the Desirability of a Culturally Pure Society: The Purpose that Measuring Muslim Ethnic Minorities' Assimilation Serve. [Unpublished Manuscript]

Turcatti, D. (2018). The Educational Experiences of Children of Moroccan Migrants in the Netherlands: Place-Making Against a Backdrop of Racism, Discrimination and Inequality. In M.

Montero-Sieburth (Ed.)., Special Issue: South to North Intercultural Dialogues: Cases and Experiences of the Second Generation. Journal of Intercultural Education. Taylor and Francis.

Turcatti, D., Berger, M., & Kruimer, R. (2017). Situating the Meanings of the Gap Year from the Perspectives of Moroccan-Dutch Youth Living in Amsterdam. [Unpublished Manuscript].

Turcatti, D. & Scott, H. (2015). Moroccan Residential Patterns in Amsterdam: Factors of Influence and Implications for Housing Policy. [Unpublished Manuscript]

Van De Werfhorst, H. G. (2011). Skills, Positional Good or Social Closure' The Role of Education across Structural'Institutional Labour Market Settings. Journal of Education and Work, 24(5), 521'548. DOI:

10.1080/13639080.2011.586994

Van De Werfhorst, H. G., & Van Teburgen, T. (2007). Ethnicity, Schooling, and Merit in the Netherlands. Ethnicities, 7, 416-444. DOI: 10.1177/1468796807080236

van der Veen, I., & Wim Meijnen, G. (2002). The Parents of Successful Secondary School Students of Turkish and Moroccan Background in The Netherlands: Parenting Practices and the Relationship with Parents. Social Behaviour and Personality, 30(3), 303-316. DOI: 10.2224/sbp.2002.30.3.303.

van Gogh, T. (2004). Submission. [Video File]. Retrieved from: https://www.youtube.com/watch'v=o1

van Mourik, K., Crone, M. R., Pels, T. V. M., & Reis, R. (2016). Parents' Beliefs about the Cause of Parenting Problems and Relevance of Parenting Support: Understanding Low Participation of Ethnic Minority and Low Socioeconomic Status Families in the Netherlands. Children and Youth Services Review, 61, 345'352. DOI: 10.1016/j.childyouth.2015.12.028

van Praag, C., & Schoorl, J. (2008). Housing and Segregation. In M. Crul & L. Heering, The Position of and Moroccan Second Generation in Amsterdam and Rotterdam. Amsterdam: Amsterdam University Press.

van Zenderen, K. L. J. (2010). Young Migrants' Transition from School to Work: Obstacles and Opportunities. [Doctoral Dissertation]. Retrieved from: http://www.verbeteronderzoek.nl/uploaded_files/regular/ProefschriftKajvanZenderenSIGRVO.pdf

Vasta, E. (2007). From Ethnic Minorities to Ethnic Majority Policy: Multiculturalism and the Shift to Assimilationism in the Netherlands. Ethnic and Racial Studies, 30(5), 713'740. DOI: 10.1080/01419870701491770

Vermeulen, L., Castellar, E. N., Janssen, D., Calvi, L., & Van Looy, J. (2016). Playing Under Threat. Examining Stereotype Threat in Female Game Players. Computer un Human Behavior, 57, 377-387. DOI: 10.1016/j.chb.2015.12.042

Verkuyten, M., & Thijs, J. (2017). Racist Victimization among Children in The Netherlands: The Effect of Ethnic Group and School. Ethnic and Racial Studies, 25(2), 310-331. DOI: 10.1080/0141987012010950

Verkuyten, M., Thijs, J., & Stevens, G. (2012).

Multiple Identities and Religious Transmission: A Study Among Moroccan-Dutch Muslim Adolescents and Their Parents. Child Development, 83(5), 1577'1590. DOI: 10.1111/j.1467-8624.2012.01794.x

Vervoort, M. H. M., & Scholte, R. H. J. (2010). Bullying and Victimization among Adolescents: The Role of Ethnicity and Ethnic Composition of School Class. Journal of Youth Adolescence, 39, 1-11. DOI: 10.1007/s10964-008-9355-y

Vigil, J. D. (2004). Gangs and Group Membership: Implications for Schooling. In M. A. Gibson, P. Gandara & J. P. Koyama (Eds.), School Connections: U.S. Mexican Youth, Peers, and School Achievement. (pp. 87-106). New York: Teachers College Press.

Weerman, F. M., Bernasco, W., Bruinsma, G. J. N., & Pauwels, L. J. R. (2015). When Is Spending Time with Peers Related to Delinquency' The Importance of Where, What and With Whom. Crime & Delinquency, 6(10), 1386-1413. DOI: 10.1177/0011128713478129

Weis, L., & Fine, M. (2012). Critical Bifocality and Circuits of Privileged: Expanding Critical Ethnographic Theory and Design. Harvard Educational Review, 82(2), 173-201. DOI: 10.17763/haer.82.2.v1jx34n441532242

Wekker, G. (2016). White Innocence: Paradoxes of Colonialism and Race. Durham and London: Duke University Press.

Welsch, W. (1999). Transculturality: The Puzzling Form of Cultures Today. In M. Featherstone & S. Lash (Eds.), Spaces of Culture: City, Nation, World (pp. 194-213). London, UK: SAGE Publications.

Wissink, I. B., Dekovic, M., & Meijer, A. M. (2009). Adolescent Friendship Relations and Developmental Outcomes: Ethnic and Gender Differences. Journal of Early Adolescence, 29(3), 405'425. DOI: 10.1177/0272431608322945

Zaal, (2014). In the Shadow of Tolerance: The Discursive Context of Dutch-Born Muslim Youth. Policy Futures in Education, 12(1), 111-123. DOI: 10.2304/pfie.2014.12.1.111

Living antiquity: the Borghese Hermaphrodite as Baroque installation

Arent Boon

Abstract

In the early seventeenth century, the Borghese Hermaphrodite was the centrepiece of an elaborate installation in the casino of the Villa Borghese, Rome. Although the discovery of this antique statue of a lying Hermaphrodite was rooted in early modern interest in Roman antique sculpture, this installation created a dynamic illusion of this artwork coming to life. This Capstone researches how the antique body of the Borghese Hermaphrodite was transformed into a Baroque installation in which beholders experienced this body's living presence. Although the restorations of the antique statue emphasised the Hermaphrodite's physical tension, Bernini's mattress more significantly provided the stage for the theatrical performance of the Hermaphrodite. In this Baroque installation, visitors would experience the living presence of the Hermaphrodite as a result of their movement in response to the sculpted illusion of movement. The environment of Scipione Borghese's villa and his casino equally provided tensions between aesthetic pleasure and immersive surprise. This practice reflects the Baroque cultural desire for the experience of antique sculpture as living bodies.

Keywords and phrases: gender, installation art, living presence, reception of antiquity, sculpture

Introduction

Most visitors will be pleasantly surprised when seeing the Borghese Hermaphrodite for the first time.[1] Their gaze rarely escapes the Hermaphrodite's graceful pose, gentle curve and lifted buttocks in the Louvre's Salle des Caryatides (figure 1). Modestly covered in drapery, the figure's ideal body lies almost asleep on an intricately carved mattress and pillow (figure 2). Its dancing left foot and leg lifted in the air engages its viewers to walk around the sculpture. On this other side, the sight of a penis appears (figure 3). Resting in between a partially exposed breast, hip, and drapery, this penis turns a beautiful body into something confrontational. The twist of this body, trapped inside white Carrara marble, suggests the Hermaphrodite reveals itself right before its return to concealment.

Following this surprise at first sight, visitors read that the sculpture dates from the 1st century AD, and the mattress from the 17th century. They learn that the statue depicts Hermaphroditus, divine son of Hermes and Aphrodite, whose body merges with the nymph Salmacis after an amorous encounter. However, these visitors to the Louvre will not be able to know how the experience of the Borghese Hermaphrodite must have been in the space for which this ancient body was first restored and this

mattress was sculpted. 17th-century visitors to the casino of the Villa Borghese would have seen this ensemble in a large, virtually empty, and specifically designated room in which the statue and mattress were installed on top of a walnut chest (cassone), unveiled after opening a lid (coperchio) on top. As the visitors would walk through this room, furnished with a mattress probably similar to the Hermaphrodite's, the opening and closing of this room's large windows would have allowed light to dance around the Hermaphrodite's seemingly moving body. The 17th-century theatrical installation sharply contrasts the aesthetic and historicist emphases of the Louvre display of the Borghese Hermaphrodite.

This Capstone researches how the antique body of the Borghese Hermaphrodite was transformed into a Baroque installation in which beholders experienced the body's living presence. An introductory section now follows with three aims. Firstly, a clearer impression will be sketched of what is known about the 17th-century Baroque installation of the Borghese Hermaphrodite. Secondly, this sculpture's state of research will be assessed. Lastly, the methodological framework this thesis employs will be clarified.

In the summer of 1619, in the gardens of the monastery of Santa Maria della Vittoria in Rome, a statue of a lying Hermaphrodite missing its nose,

---

[1]The author would like to thank this thesis supervisor Minou Schraven, Harald Hendrix, Arno Witte, Joris van Gastel, and Janna Schoenberger, as well as the Royal Dutch Institute (KNIR) in Rome, the Dutch Institute for Art History (NIKI) in Florence and the Rijksmuseum for their hospitality

[2]This follows the hypothesis raised of the Hermaphrodite's discovery as proposed in Kalveram, Scipione Borghese, 121. Sketches of a body resembling the lying Hermaphrodite attributed to Rubens in the British Museum and Metropolitan Museum of Art have led to the now disproven assumption the excavation could have occured in 1609. See Haskell and Penny, Taste and the antique, 234-236. Kalveram's assumption that the statue has been found in the garden has been debated. See Roani Villani, 'L'ermafrodito',

some fingertips, a left foot and its base was discovered[2]. Following its excavation, the statue of the Hermaphrodite was acquired from the frati for a substantial price by cardinal Scipione Borghese.[3] The Borghese family, originally from Siena, migrated to Rome in the late 16th century. Following the election of his uncle Camillo Borghese to the papacy as Paul V in 1605, Scipione was appointed cardinal-nephew.[4] This position effectively entailed the responsibility for the collection and exhibition of Roman antique sculptures as a justification for this political ascent. Indeed, Scipione swiftly concerned himself with the mass acquisition of collections of antique sculptures.[5] A series of payments in the Borghese archives, now preserved in the Vatican, provides detailed insight into how a fragmented statue of a Hermaphrodite eventually turned into an installation in the casino of the Villa Borghese. It may have been first sent to the workshop of Giovanni van Santen, the Borgheses' in-house architect, where his assistant, David Larique, conducted its restoration.[6] At the same time, the workshop of emerging artist Gianlorenzo Bernini would have received marble to eventually carve the hermaphrodite's mattress.[7] Subsequently, the architect Giovanni Battista Soria carved the cassone and the coperchio.[8]

The first documentation of the installation of the Borghese Hermaphrodite could be traced to were travel guides dating from the late 1630s onwards.[9] The Hermaphrodite and Bernini's mattress enjoyed equal fame in the early modern period.[10] Following the moving of the statue to a more constrained space on the ground floor of the Galleria Borghese,

the Borghese Hermaphrodite was sold, like many other antiquities from the Borghese collection, to Napoleon by Camillo Borghese in 1807. In the exhibition space of the Louvre's department of Greek and Roman antiquities, its former Baroque installation remained out of sight and out of mind. In the mid-1990s, a series of exhibitions on Bernini and the Borghese collection of antiquities brought a more positive kind of attention to the statue.[11] At this same time, archival research illuminated the historical context in which the hermaphrodite was acquired and converted into an installation.[12] The Borghese Hermaphrodite has rewarded attention from multiple art historical disciplines, such as the reception of antiquity, the history of collecting, the history of sculpture, and museology. However, in the present state of research, two arguably unsatisfying tendencies may still be identified. Firstly, art historians are inclined to categorise the statue-mattress ensemble as either an example of Roman antique sculpture or Bernini's restoration practice. The Borghese Hermaphrodite fulfils neither expectation. Moreover, as the Borghese Hermaphrodite's Baroque installation has largely escaped analysis, aesthetic or semiotic approaches remain most frequently applied to this case study.

This Capstone aims instead to understand the social context in which the Baroque installation of the Borghese Hermaphrodite was experienced. This requires a theoretical framework that exceeds traditional art historical disciplinary boundaries. The anthropologist Alfred Gell, in his Art and Agency (1998), famously defined the art nexus as a social framework in which art works operate

---

40.

[3]As testified by a documented payment from the Borghese Archives in the Vatican dated June 26, 1619. See Coliva, eds., Bernini scultore, Appendix, 430 - 431.

[4]For the political and economic status of Scipione Borghese, see Reinhardt, Kardinal Scipione Borghese (1984).

[5]The Borghese acquired in 1607 the collection of the bankrupt bankers Ceuli and in 1609 from Tommaso della Porta's family of sculptors. See Kalveram, Scipione Borghese, 7 ' 16.

[6]As testified by a payment of 15 scudi made to Larique dated February 19, 1620. For the source of these payments, see note 2.

[7]As testified by a payment of 4 scudi and 30 baiocchi made to the facchini dated September 29, 1619, a payment of 15 scudi to Bernini for marble acquisiton dated October 14, and a payment of 60 scudi to Bernini for the restoration of the statue and the creation of the mattress dated February 21, 1620.

[8]As testified by a payment of 230 scudi made to Soria.

[9]For early modern sources, see mainly Perrier, Segmenta nobilium et signorum (1638), Martinelli, Roma ricercata (1644), Manilli, Villa Borghese (1650), Montelatici, Villa Borghese (1700), Raguenet, Les monumens de Rome (1701), Maffei, Raccolta di statue antiche e moderne (1704), Brigentius, Villa Burghesia (1714).

[10]Haskell and Penny, Taste and the antique, 234 - 236.

[11]The Borghese Hermaphrodite returned from the Louvre to the Galleria Borghese for Bernini scultore (1997), I Borghese e l'antico (2011) and Bernini (2017).

[12]See Kalveram, Scipione Borghese (1995), Roani Villani, 'L'ermafrodito' (1993) and note 11.

[13]See Gell, Art and agency (1998).

[14]Gell, Art and agency, 16 - 17.

as agents of human behaviour and vice versa.[13] Agency can be attributed to persons and things when initiating causal sequences of behaviour.[14] In this art nexus, Gell distinguishes the index, prototype, artist, and recipient. The index, a material entity that motivates inferences, responses or interpretations, may be equalled to the art work.[15] The index functions as a representation of the prototype when there is resemblance, triggering recognition.[16] The artist is the agent of manufacture to whom the origin of the index is attributed.[17] The spectator to the art work operates as the recipient in the art nexus, mostly as patient, but also as agent, in the case of a patron.[18]

Alfred Gell understands the relationship between object and subject, or index and recipient in his own terms, as proceeding from agent to patient.[19] This premise, however, rewards complication. Caroline van Eck, in Art, Agency and Living Presence (2015) proposes that the recipient of the index should be framed as an agent in order for their experience to be understood.[20] She conceives the concept of living presence to describe the beholder's experience of a work of art coming to life before the eyes of this spectator.[21] This thesis will follow the premise that such a response, crossing the boundary between object and subject, occurs when an art work actively manages to respond to the world of the beholder whilst giving shape to this world.[22] According to Van Eck, the living presence response temporarily destroys the aesthetic or historical significance of an art work, making spectators forget they are looking at an image.[23] Following this premise, this Capstone explores precisely how in the experience of the Baroque installation of the Borghese Hermaphrodite, the historical significance of the antique body was disrupted.

This research question results in four different chapters. Each is dedicated to a component of the transformation of an archaeological artifact into a theatrical installation. The first, "Discovery", explores how the Hermaphrodite's gender defined the historical significance of the discov-

ered antique body. The second chapter, "Restoration", assesses how these restorations conducted on this archaeological fragment respected its antiquity. "Performance", the third chapter, considers how the presence of Bernini's mattress challenges the historical significance of the antique body. The fourth and final chapter, "Installation", researches how the Baroque experience of the Borghese Hermaphrodite was mediated by the spaces in which it was installed.
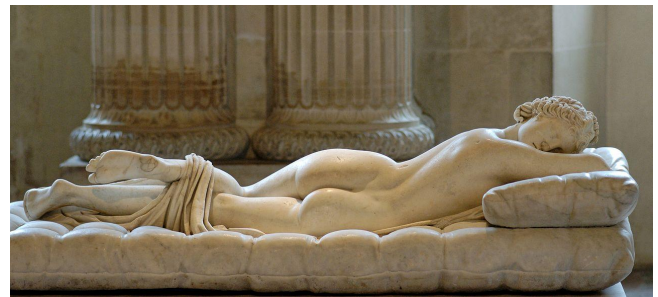
Borghese Hermaphrodite, c. 1st-2nd century AD, including additions by Gianlorenzo Bernini, c. 1619 ' 1620. Marble copy after bronze model, Musée du Louvre, Paris. Courtesy of Wikimedia Commons.)

Borghese Hermaphrodite, seen from back. Courtesy of Wikimedia Commons.)

---

[15]Ibidem, ix, and 12 - 16.

[16]Ibidem, 25 - 26.

[17]Ibidem, 23.

[18]Ibidem, 24.

[19]See Gell, 'The technology of enchantment and the enchantment of technology' (1992).

[20]Van Eck, Art, agency and living presence, 52 ' 53.

[21]Van Eck, Art, agency and living presence (2015) and 'Living statues', esp. 644 - 650.

[22]Van Gastel, Il marmo spirante, 1 ' 5, 38 ' 40, 207.

[23]Van Eck, Art, agency and living presence, 144 - 157. See Nagel and Wood, Anachronic Renaissance (2011).

Borghese Hermaphrodite, seen from back.
Courtesy of Wikimedia Commons.)

I Discovery

Around the time of the discovery of the Borghese Hermaphrodite, another cardinal, Francesco Maria del Monte, was equally occupied with the excavations around Santa Maria della Vittoria. In a letter to a friend dated July 5, 1619, he describes at length the sculpture found in the garden of this monastery, which he describes as 'a most wonderful statue of a woman who awakens as a man, but sir Cardinal Borghese has wanted it'.[24] This reflects how the Hermaphrodite's antiquity was already desired when it was first discovered. This chapter explores how the Borghese Hermaphrodite's gender defined the historical significance of the discovered antique body.

The mythical birth of the Hermaphrodite originates with Ovid's Metamorphoses. In the surroundings of an idyllic lake, the nymph Salmacis encounters Hermaphroditus, son of the gods Hermes and Aphrodite.[25] Salmacis falls rapturously in love with him. She prays to the gods for their two bodies to be united. Her prayers were answered:

> "[F]or their two bodies, joined together
> as they were, were merged in one, with

one face and form for both. As when one grafts a twig on some tree, he sees the branches grow one, and with common life come to maturity, so were these two bodies knit in close embrace: they were no longer two, nor such as to be called, one, woman, and one, man. They seemed neither, and yet both."[26]

In this myth of origin of the Hermaphrodite, three aspects stand out, namely this deity's gender, the erotically charged circumstances that created this body, and the natural, ideal surroundings in which this metamorphosis occurs. In Ovid's narrative, Hermaphroditus receives divine status by virtue of its gender, neither fully man, nor fully woman. Therein lies the gendered confrontation essential to the cultural locus of the Hermaphrodite.

The Hermaphrodite emerges in Neoplatonic and Aristotelian philosophical debates on the origin of the human species and the gender binary.[27] In Roman antiquity, the Hermaphrodite was integrated into its polytheistic religious culture, representing love, lust and the harmonious unification of genders.[28] Three sculptural types were consequently developed, namely the standing, struggling and sleeping Hermaphrodite.[29] The sleeping Hermaphrodite gained popularity in early imperial Rome and was generally exhibited in bathhouses or gardens of villae.[30] This type, of which the Borghese Hermaphrodite qualifies as the foremost example, consists of the body sleeping on a rock, partially covered in drapery, the hip lifted to reveal the penis.[31] In the Palazzo Massimo, Rome, a largely unrestored example of this lying type is preserved, creating an impression of how the newly excavated Hermaphrodite may have looked (figure 4).[32] Thus, the sculptural model of the lying Hermaphrodite locates this gendered body in natural surroundings and thus echoes the spatial contexts proposed in

---

[24] Cited in Roani Villani, 'L'ermafrodito', 45 - 46.

[25] See Ovid, Metamorphoses, Book IV, 274 - 388.

[26] Ovid, Metamorphoses, Book IV, 372 - 379.

[27] Long, Hermaphrodites, 7 - 14.

[28] For the cultural locus of the Hermaphrodite in Graeco-Roman antiquity, see Ajootian, Natus diformis (1990, unpublished). For recent insights, see Brisson, eds., Deep classics (2016) and Von Stackelberg, 'Garden hybrids' (2014).

[29] Von Stackelberg, 'Garden hybrids', 399 and note 10, also for a literature review

[30] Ibidem, 399.

[31] Pliny the Elder mentions in his Natural history that 'Polycles made a famous (sublimis) Hermaphrodite' (Book XXXIV, 80 ' 81). See Mancinotti, Ermafrodito dormienti (2018).

[32] This Palazzo Massimo Hermaphrodite was retrieved in 1879 in Rome during the construction of the Teatro Costanzi and entered this collection before 1882.

the Ovidian narrative.

In the early modern period, the desire for the discovery of antique objects such as the sleeping Hermaphrodite rapidly developed. In this epistemological shift, the Hermaphrodite's gendered confrontation was complicated by its historical significance.[33] The Sicilian humanist Antonio Beccadelli explores this nexus between gender, sexuality and antiquity in Hermaphrodite (1425 - 1426). The Hermaphrodite is a book of epigrams that, following its Roman pursuer Martial, juxtaposes blunt erotic humour with references to antique gods. Beccadelli explicitly states his book is divided into two parts, like a Hermaphrodite, the first 'stand[ing] for the cock', the second for the 'cunt'.[34] In his introductory epigram, titled 'That he may [...] read this book sympathetically, even though it is naughty, and so join the author in imitating the ancients', he announces that

> '[T]his work rouses laughter from any-
> one, no matter how stiff and serious
> and stirs anyone's loins, even Hippoly-
> tus'.
> And in this I follow the example of the
> learned poets of old,
> who it is clear, composed trifle
> and, it is evident, lived modest lives,
> even if their pages were full of obscene
> jokes.
> The lazy crow fails to notice this, who
> have no care to look to the ancients
> but whose only care has been given to
> their belly.
> their ignorance will pick at my trifles
> too:
> so be it ' the learned will not reproach
> me".[35]

Beccadelli's Hermaphrodite provides insight into a cultural field of tension between the Hermaphrodite's gender and its antiquity. On the one hand, its body has the power to 'stir anyone's loins', its gender charged with erotic potential. On the other hand, Beccadelli, perhaps in spite of his erotic fascination for the Hermaphrodite, explicitly locates himself amongst 'the learned poets of old'. He justifies the Hermaphrodite's gendered confrontation by underlying its historical significance. This humanist interest in the Hermaphrodite attempts to address the gendered confrontation by emphasising its antique origins.

Nevertheless, the potential of direct, emotional affect remains when encountering the Hermaphrodite. The Florentine artist Lorenzo Ghiberti, contemporary of Beccadelli, saw an example of the sleeping sculptural type. In his Commentaries (c. 1447), he describes his experience of

> "A Hermaphrodite, the size of a thirteen-
> year-old girl, which had been made with
> admirable skill [...] It is impossible for
> the tongue to tell the perfection and the
> knowledge, art and skill of that statue.
> The figure was on spaded earth over
> which a linen cloth was thrown. The
> figure was on the linen cloth and was
> turned in a way to show both the mas-
> culine and the feminine characteristics
> [...] one of the legs was stretched out
> and the large toe had caught the cloth,
> and the pulling of the cloth was shown
> with wonderful skill. [...] The eye per-
> ceived nothing if the hand had not found
> it by touch."[36]

Ghiberti articulates how the Hermaphrodite's gender characterises the historical significance of this archaeological object.[37] By this time, the archaeological recovery of antique sculptures was a long established practice.[38] Antique statues of Hermaphrodites had already been discovered and put on display in Roman collections throughout the fifteenth and sixteenth centuries.[39] Ovid's story of Hermaphroditus and Salmacis remained well known throughout the early modern period

---

[33] See Barkan, Unearthing the past (1999).

[34] Beccadelli, Hermaphrodite, Book I, XLII.

[35] Ibidem, Book I, I, 3 ' 10.

[36] As cited in Barkan, Unearthing the past, 164 - 165. Translated by Elizabeth Holt.

[37] Leonard Barkan argues that Ghiberti here accounts for the mimetic confusion the Hermaphrodite's body entails. See Unearthing the past, 158 - 166.

[38] For Renaissance archaeology, see Wren Christian, Empire without end (2010), and Furlotti, Antiquities in motion (forthcoming).

[39] The Altieri in Rome owned a statue of a standing Hermaphrodite, to be later covered with drapery, and now in the Palazzo Doria Pamphilj. See Wren Christian, Empire without end, 258 ' 259.

[40] In Edmund Spenser's Fairie Queene (1590), two lovers 'seemd [...] growne together quite' like 'that faire Hermaphrodite, which

via translations and revised editions.[40] Ghiberti must then have been one of many to admire the Hermaphrodite for its antiquity. However, its gender demanded to Ghiberti and to many others an immediate physical interaction that transcended this body's history. When the Borghese Hermaphrodite was discovered, as a feminine body revealing its masculinity, a historical artifact asserts itself into the present day.

This chapter's quest for the understanding of the Hermaphrodite's gender in its early modern reception, several premises have emerged. The Roman antique sculptural model of the sleeping Hermaphrodite creates a gendered confrontation in the form of a female body with masculine genitalia. In the early modern reception of the Hermaphrodite, this gendered confrontation is complicated by its perceived historical significance. The humanist interest in the Hermaphrodite is characterised by the acknowledgement of its erotic potential insofar that this potential can be justified by the Hermaphrodite's antique origins. As the next chapters will show, in the Baroque life of the Borghese Hermaphrodite, this statue's antique past will continue to be challenged.

Sleeping Hermaphrodite, mid-2nd century AD. Marble, Palazzo Massimo, Museo Nazionale Romano, Rome. Image courtesy of Metropolitan Museum of Art, New York.)

II Restoration

When the Borghese Hermaphrodite was unearthed, it asserted itself into the present of its viewers. When this statue was restored, its antique body was charged with the potential to seem alive. David Larique, workshop assistant to Giovanni van Santen, the Borghese's foremost architect, conducted the restoration of the Borghese Hermaphrodite, probably in late 1619 or early 1620.[41] This chapter explores how his restorations to the Borghese Hermaphrodite respected the antiquity of the unearthed statue.

It is reasonable to presume restorations recreate a past reality of a historical object. However, the process is more complex. Paradoxically, restorers impose their artistic touch onto historical objects as a valuation of their historicity. The early 17th century in Rome is often marked for a shifting trend from the antiquarian admiration of sculptural fragments to the desire of the return of fragments to complete bodies.[42] An artistic tension then emerges between the unearthed historical object and its Baroque restorations.[43] Restorers of antique sculpture in the early modern period were often aspiring artists themselves. The sculptor Orfeo Boselli, in his unpublished treatise Osservazioni sulla scultura antica (c. 1650) argues the restorer should aim to 'accompany the antique manner'.[44] This 'good' restorer, in order to 'know' the antique statue, must identify the 'Deity', so that the statue can be given the 'suitable symbols' and the 'proper proportions'.[45] In a chapter on the 'Methods of Restoring', Boselli describes how to add limbs by creating holes in the original marble in which the newly carved piece could be cemented and covered with plaster as to prevent the cracks from being visible.[46] According to Boselli, the restorer's accompaniment of antique sculpture equals its partial physical destruction with the ultimate aim of turning this fragment into a complete body. From

---

that rich Romane of white marble wrought, and in his costly Bath causd to bee site' (Book III, canto XII, stanza XLVI). For the early modern literary reception of Ovid's Metamorphoses, see Stanivukovic, eds., Ovid and the Renaissance body (2001).

[41]See note 2 and 5

[42]Barkan distinguishes imaginative from discursive restoration in favour of this narratives of development. See Unearthing the past, 210. Haskell was famously critical of 'ruthless and imaginative' restorations. See Taste and the antique, 25 - 27. For the Borghese's collectionism and restoration, see Coliva, 'La storia di un famiglia', in Coliva, eds., I Borghese e l'antico, 19 - 45.

[43]Montagu, 'The influence of the Baroque on classical antiquity', in Beck and Schütze, Antikenrezeption im Hockbarock, esp. 95 - 97, and Sparti, 'La 'rinascita' di sculture antiche', in Bacchi and Coliva, eds., Bernini, 75 ' 83.

[44]Cited in Dent Weil, 'Contributions', 86 - 88. See also Kalveram, Scipione Borghese, 90 ' 92. For a social reading of Boselli, see Fortunati, 'Boselli', in Colantuono and Ostrow, eds., Critical perspectives, 105 - 116

[45]Cited in Dent Weil, 'Contributions', 86 ' 88.

[46]Ibidem, 90 - 95.

Boselli's treatise, a paradox emerges. Restorers are most successful when their sculptural virtuosity blends in with the actualised historical object, to the point of no longer being distinguishable.

The cracks and varying shades of white of this marble luckily render the restorations to the body of the Borghese Hermaphrodite visible (figure 5). The careful observer may now notice that the nose, hands and the right foot are probably not a part of the original antique statue.[47] Presumably, then, these fragments are a part of Larique's restorations. Larique must have followed the methodology Boselli would later record. He surely identified the Hermaphrodite's iconography, determined its original pose, and added pieces to convert this fragmented body into a complete whole. The question pertains here how Larique 'accompanied' the Borghese Hermaphrodite's antiquity. A crack above the ankle of the statue's left foot reveals it must have been added (figure 6).[48] This piece of marble has been carved in such a manner its muscles seem stretched, its toes dancing in the air, following the dynamism of the legs, lifted buttocks, and curve around the back. Effectively, a 17th-century marble foot is attached to an antique body, with the result of emphasising this body's tense physical movement. For his restorations of the Borghese Hermaphrodite, David Larique accompanied antiquity in the sense that his sculptural assertions of artistic presence served to emphasise the physical tensions the antique statue confronted him with.

A second example of the sculptural type of the lying Hermaphrodite is at present displayed in the Uffizi (figure 7). Ludovisi Hermaphrodite's social life was lived in this same era. Excavated between 1621 and 1623, the sculpture entered the collection of cardinal Ludovico Ludovisi.[49] Ludovisi was appointed by his uncle, Pope Gregory XV, who was elected following the death of Paul V Borghese in 1621. Like Scipione Borghese, Ludovico established a villa around the Porta Pinciana in Rome. Like Scipione, Ludovico here ordered the construction of a pavillon, presently known as the Casino dell'Aurora, for the display of his collection of paintings and sculptures.[50] However, the Ludovisi, as the successive papal family to the Borghese, would have developed a system of patronage to distinguish them from their predecessors.

Ludovico Ludovisi appointed the sculptor Ippolito Buzzi as the restorer of his collection of antiquities.[51] Buzzi's accompaniment of antiquity notably equalled the habitual conversion of torsos and limbs into complete bodies.[52] In the case of the Ludovisi Hermaphrodite, its body, although without legs, and its head are probably antique.[53] Buzzi's restorations of the legs and feet, however, notably differ from its predecessor (figure 8).[54] Buzzi's feet lie completely relaxed on the rock, the carving of its toes arguably lacking the intricacy of their Borghese counterpart. As a result of these restorations, the body of the Ludovisi Hermaphrodite suggests relaxation to a greater extent than the Borghese's. Whereas Buzzi's restorations follow the stillness as directed by the statue and, relevantly, the antique marble base in the shape of a rock, Larique's restorations follow the potential for movement provided by the antique fragment.

The Borghese and Ludovisi restorations reflect two technically convergent yet practically divergent approaches to the accompaniment of antiquity in the early 17th-century. For both David Larique and Ippolito Buzzi, the desire for the

---

[47]Kalveram, Scipione Borghese, 120. It is relevant to emphasise that Bernini's mattress, to be discussed in Chapter III, is often mistakenly understood as an example of the restoration of the Borghese Hermaphrodite. See note 2, 5 and 42.

[48]The Hermaphrodite's foot may still be attributed to Bernini instead of Larique. Bernini has also been paid for the statue's restorations. See note 5 and 6.

[49]The Ludovisi Hermaphrodite has not acquired the renown of its Borghese predecessor. See Bruand, 'La restauration', 410, and Haskell and Penny, Taste and the antique, 235 - 236

[50]The Ludovisi Hermaphrodite was probably displayed here. See Amadio, 'La villa Ludovisi', in Giuliano, eds., La collezione Buoncompagni Ludovisi, 9 - 17, and Felici, Villa Borghese (1952). The Ludovisi Hermaphrodite was acquired for an extraordinary price by Ferdinand II de' Medici in 1669, after which the statue was moved to Florence. See Haskell and Penny, Taste and the antique, 235 ' 236.

[51]For the restoration of the Ludovisi collection of antiquities, see Bruand, 'La restauration' (1956) and Marvin, 'Possessions of princes', in Grossman, Podany and True, eds., History of restoration, 225 - 238.

[52]Bruand, 'La restauration', esp. 397 - 398. A notable example of Buzzi's restorations for the Ludovisi includes a fragmented statue of Salmacis embracing Hermaphroditus restored into Amor and Psyche, currently displayed in the Palazzo Altemps, Museo Nazionale Romano.

[53]The antiquity of the base in the form of a rock remains a subject of debate.

[54]Interestingly, the assumption has pertained that Buzzi's restorations were modelled after Larique's. See Haskell and Penny, Taste and the antique, 235 ' 236.
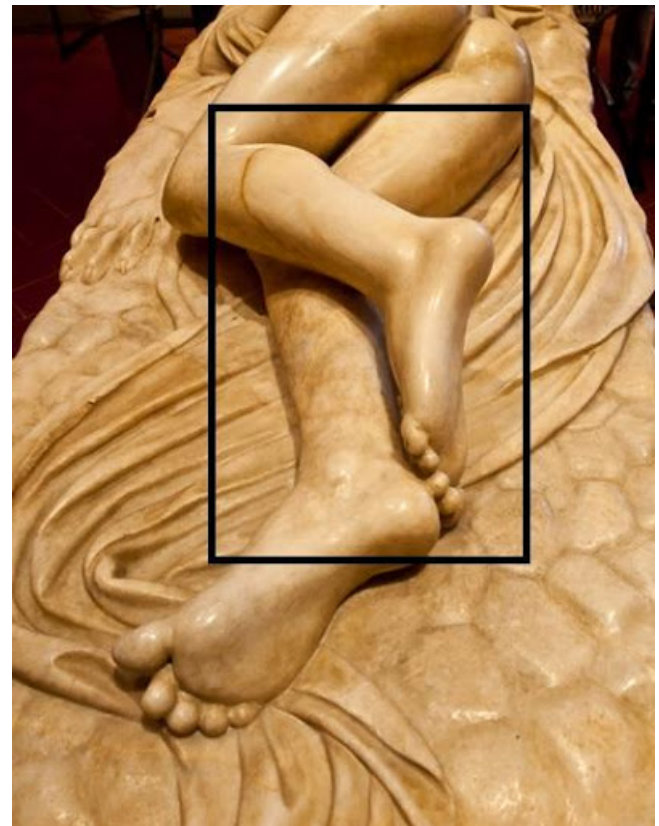
restoration of an object results from the admiration of the object's historical significance, in spite of these restorers' paradoxical imposition of their own sculptural presence onto these antique objects. Technically, in both cases, the implicit tension between antiquity and modernity is explicitly denied. In a physical sense, both sculptors presume the fragmented excavated object as a point of departure towards its completion. However, their respective executions vary significantly. The Ludovisi Hermaphrodite's legs suggest the body as still and resting. The Borghese body, on the contrary, suggests physical tension and movement as a result of its restoration. The restorations to the Borghese Hermaphrodite respected the antique statue in the sense that they charged the antique fragment with the potential to animate. It is this capacity to stage living presence on which Bernini's artistic involvements rely, as is shown in the next chapter.

Lying Hermaphrodite, known as Ludovisi Hermaphrodite, c. 1st-2nd century AD. Marble copy after bronze model [...] , Galleria degli Uffizi, Florence. Installation view. Courtesy of Flickr.)
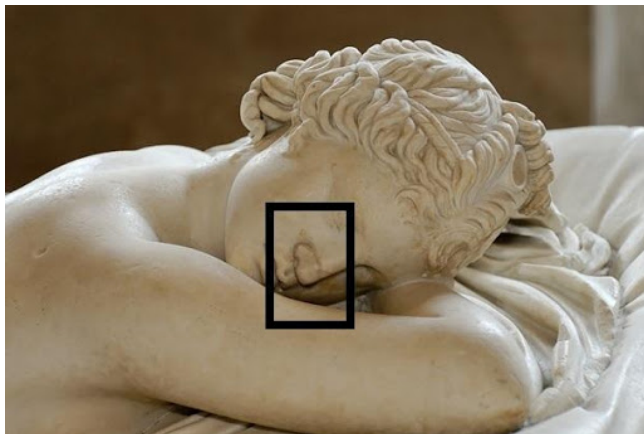
Borghese Hermaphrodite, detail of head. Courtesy of Wikimedia Commons. Restoration indicated.

Borghese Hermaphrodite, detail of back, with feet. Courtesy of Flickr. Restoration indicated.)

Ludovisi Hermaphrodite, detail of back, with feet. Courtesy of Flickr. Restoration indicated.)
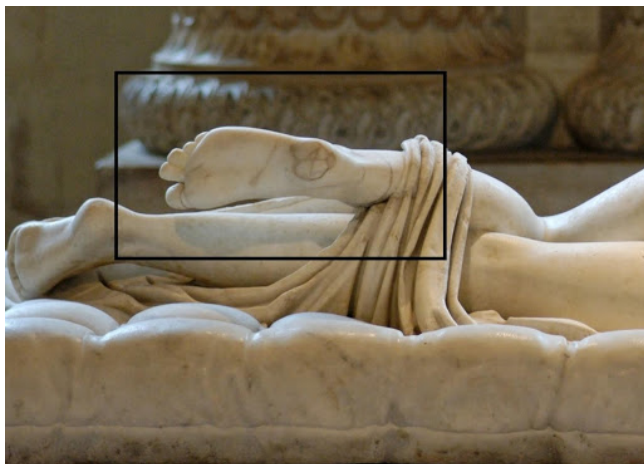
III Performance

In 1701, the French abbot François Raguenet published a guide, Les monumens de Rome, in which he describes the Borghese Hermaphrodite at length.[55]   Following his aim to 'present to

---

[55]Raguenet, Les monumens de Rome, 31 ' 33. Translated by author. It is probable Scipione's installation no longer fully existed at this time.

the imagination' of his readers 'the rarest creations [...] in the Arts, brought by [...] the Romans', Raguenet is attracted the most to Bernini's mattress.[56]   He enthusiastically describes how '[a]nyone [...] lays their finger on [the mattress], and everyone senses, with an unknown kind of trembling horror, the hardness of the resisting marble even though it was normal to believe the finger would sink.'[57]   Raguenet's account describes the confusion between materiality and immateriality that Bernini's artistry incites.[58] His anecdote raises questions about Bernini's artistic agency over the spectator's experience of the installation of the Borghese Hermaphrodite. This chapter will explore how the mattress' material confusion, incited by Bernini's artistic presence, interacted with the antique statue's historical significance.

The living presence of art, or the possibility of an emotional and impulsive experience of an artistic object as if it is not created, but actually real, constituted a major source of intellectual debate in the early 17th-century.[59] Pope Alexander VII's rejection of a monument offered to him by the Roman people, for example, resulted in a debate surrounding the relationship between a statue and the living being it represents.[60] In 1644, the Jesuit philosopher Sforza Pallavicino published Book IV of Del bene, a treatise that explores the psychological experience of an inanimate object as if it were a living thing.[61] In this treatise, Pallavicino defines what he calls the prima apprensione, a first, pre-cognitive interaction with an object before the rational truth to this experience is defined. This prima apprensione awakens the fantasia, or the memories of reality which stimulate the belief in an artwork as a real being. Pallavicino finds that

"[T]o imitate [...] means to produce with one's work some of the sensible effects [...] which one usually finds only in the object being imitated; while, if it occurs that the same effects are found elsewhere, promptly, they will awaken in the imagination the memory of that object in which it is most commonly found, and of the other of its properties that we were used to experience".[62]

Following Pallavicino, the 'sensible effects' of artworks awaken the memory of the represented object and of the properties that instigate an experience of this object.[63] The experience of an inanimate artwork as something real accordingly originates with the object's effect that activates the subject's memory of the real thing.

The art of Gianlorenzo Bernini encapsulates this cultural interest in experiences of living presence. Irving Lavin has defined Bernini's 'unitarian vision' as his capacity to create, in his artwork, 'sudden thrusts' that invade the minds of his spectators.[64] It is then productive to frame Bernini here as an agent of the technological enchantment of his viewers, invading the reality of those that get to look at it.[65]

Bernini must have received the commission from Scipione Borghese to carve a mattress for the retrieved Hermaphrodite around the time of its discovery in mid-1619.[66] In this same era, he carved a screaming head that became known as Anima dannata (figure 9).[67] This face's wrinkled forehead, large eyes and opened mouth suggest an expression in between fear and anger, resembling the face of Laocoön. Its heavy curls and stretched muscles in the neck create the illusion of movement. The head, leaning heavily forward on a small base,

---

[56]Ibidem, 'Preface'.

[57]Ibidem, 33. Raguenet attributes the restoration of the Hermaphrodite to a certain 'Guillaume Bertelot', necessarily deemed 'French by nationhood' (31).

[58]Raguenet's guide appears around the time of two biographies of Bernini's life, namely Filippo Baldinucci's (1682) and Domenico Bernini's (1713).

[59]Van Eck, Art, agency and living presence, 61 ' 99.

[60]Most specifically described in Giovanni Andrea Borboni's Delle statue (1661). See Van Eck, Art, agency and living presence, 79 ' 80.

[61]Pallavicino, Del bene, Book IV (1644). See Delbeke, The art of religion (2012), Van Eck, Art, agency and living presence, 67 - 73, and Van Gastel, Il marmo spirante (2011).

[62]Pallavicino, Del bene, Book IV, 219.

[63]Van Gastel, Il marmo spirante, 2 - 4.

[64]Lavin, Bernini and the unity, xiii and 146 - 157.

[65]See Gell, 'The technology of enchantment and the enchantment of technology' (1997). For recent insights into Bernini's (im)materiality, see Levy and Mangone, eds., Material Bernini (2016).

[66]See note 6.

[67]For a recent discussion of Anima dannata, see Cueto, 'Original meanings' (2015).

threatens to lose its stability, not just as an object, but as a marble fragment approaching human reality. Bernini's sculpture of Anima dannata expresses his artistic capacity to elicit responses to carved, white Carrara marble that echo the human experience of the real. In 17th-century rhetorical theory, the concept of contrapposti, two separate counterpoises, is used to define the interaction of opposites, or the interdependence of physical appearances to the achievement of unity of an object.[68] Bernini's Anima dannata thus achieves the unity of movement and stillness by threatening to dissolve from artwork into reality.

Bernini's mattress stages a similar kind of material confusion (figure 10). The mattress consists of flower-shaped buttons that hold the cushions together. Like the pillow lying upon the mattress, its irregular folds of the illusionary silk remain visible. In its 17th-century exhibition, another mattress, probably of similar shape and material, would have stood in the same room for viewers to contemplate the mattress.[69] When Raguenet described viewers' sense of 'trembling horror', he in fact pointed towards the 'sensible effect' that awakens in the memories of viewers the experience of a real mattress when touching Bernini's sculpture.

A closer look at the mattress and pillow shows that both follow the shape of the Hermaphrodite's body. Bernini thus implied the physical if not intellectual weight of the Hermaphrodite as an archaeological artefact. An artist like Bernini would indeed have been dedicated to studying the antique sculpture that was excavated and displayed in Rome.[70] However, the mattress bears no iconographic relationship to the Ovidian metamorphosis of Salmacis.[71] Unlike the Ludovisi Hermaphrodite's steadfast rock, Bernini's mattress underlies the movement of the twisted body that lies on top.

The previous chapter has already shown how the restorations emphasises the dynamic potential of the Hermaphrodite's body to conceal its most surprising aspect. It is this aspect of the relationship between the Borghese Hermaphrodite and its viewers that Bernini's mattress emphasises. Bernini directs the spectator's gaze in the direction of a valuable archaeological artefact. Moreover, Bernini's animate mattress and cushion charge the body that lies upon this with the potential to come to life before the eyes of its viewers. Bernini's artistry, although originating with admiration of its archaeology, enables an experience of the Hermaphrodite in which its historicity implodes.[72]

Bernini's mattress would eventually reach equal fame to the actual statue of the Hermaphrodite.[73] Travel guides to Rome and to the Villa Borghese mention both the antique Hermaphrodite and Bernini's mattress.[74] Grand Tourists often compared in their travel diaries the respective artistic complexities of the body and the mattress, or expressed their confusion about their juxtaposition.[75] Early modern sculptural reproductions equally reflect on the dynamic between this body and the mattress on which it rests. The Metropolitan Museum of Art conserves an early 1639 bronze copy of the Borghese Hermaphrodite courtesy of the Florentine sculptor Giovanni Francesco Susini (figure 11).[76] Susini has considerably recreated the aesthetic effect of the Hermaphrodite's twist, including its stretched left foot and penis. However, Bernini's mattress and pillow are considerably smaller in proportion to the original and have been carved arguably lacking the original's intricacy.[77] Most interestingly, Susini's copy of the Borghese Hermaphrodite features an inscription. The front side reads

"Duplicem formam

---

[68]Lavin, Bernini and the unity, 10 - 13. See also Summers, 'Contrapposto' (1977).

[69]Kalveram, Scipione Borghese, 74. See also Chapter IV.

[70]See Warwick, Bernini: art as theatre, 79 - 90.

[71]See Chapter I.

[72]See Warwick, Bernini: art as theatre, 79 - 119.

[73]Haskell and Penny, Taste and the antique, 234 ' 236, and note X.

[74]Manilli, Villa Borghese, 98 ' 99, Montelatici, Villa Borghese, 276 ' 277. Maffei finds the mattress' 'considerable artifice' stimulates comparisons to the antique body (Raccolta, 71).

[75]For a bibliography of tourist anecdotes of the Borghese Hermaphrodite, see Haskell and Penny, Taste and the antique, 235. Smollett notably observed the mattress' view "[...] rival[led] the softness of wool" and found it "the most curious circumstance of this article" (Travels, Letter XXXI, 263).

[76]Beyond the scope of this research lies the question of materiality of such bronze copies. It leads to wondering whether bronze, besides its material value, would influence the historicity of such copies.

[77]Susini's copy also provides a rare example of a reproduction of the cassone on which the Borghese Hermaphrodite would have stood in the early 17th century. See Chapter IV.

Uno in corpore vides
Mirare pulchritudinem'[78]

When turning the statue around, the inscription continues:

"Duplex cor uno in pectore
Saepe invenies
Cave insidias"[79]

Such an inscription may have been a part of the Borghese Hermaphrodite's installation. Bernini's Apollo and Daphne, displayed in the same casino, indeed featured such a short poem serving as a moral instruction to practices of viewership (figure 12).[80] This inscription indicates the Hermaphrodite's 'double form' necessitates aesthetic admiration, yet warns its viewers of this statue's treacherous potential.[81]

Susini's copy of the Borghese Hermaphrodite thus cautions its viewers of the sculpted body's threat to be experienced as real. Surely, Bernini's mattress serves to carry the weight of the Hermaphrodite's antiquity. More notably, its material confusion charges the antique body of the Hermaphrodite with the potential to be experienced as a living being. The artist's agency here demands a kind of emotional interaction from its recipients that challenges the statue's antiquity. Effectively, Bernini's artistic contributions result in what could be understood as a stage set to Hermaphrodite's performance. In the next and final chapter, the Borghese Hermaphrodite will be contextualised from the perspective of its theatres, so to speak, or the spaces constructed around its performance.
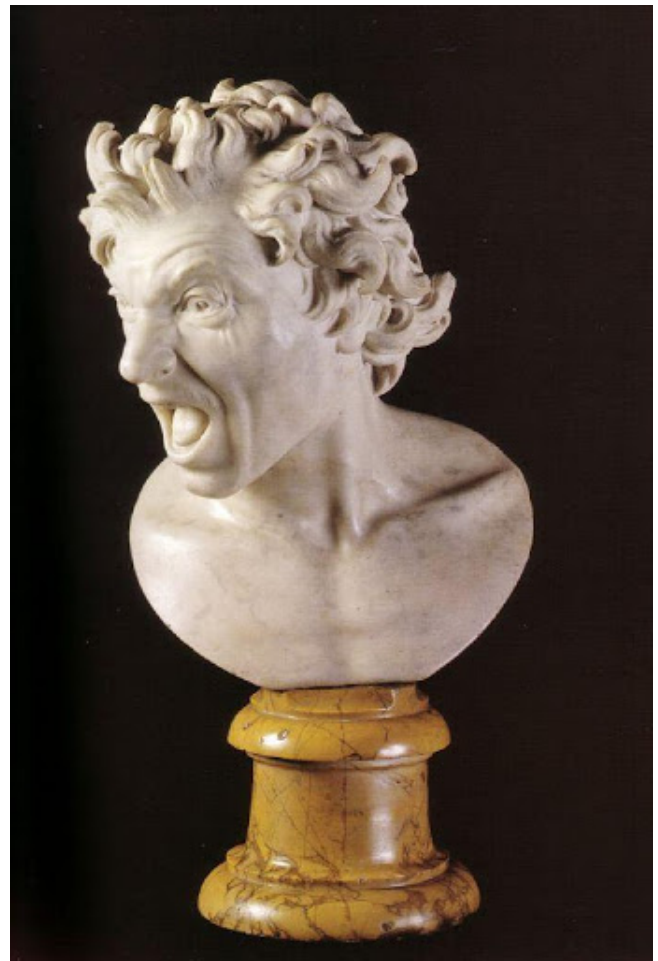
Figure 9. Gianlorenzo Bernini, Anima dannata (Damned soul), c. 1619. Marble, Palazzo di Spagna, Rome.)

Borghese Hermaphrodite, head view of body, mattress and pillow. Image retrieved from Wikimedia Commons.)

---

[78]'A double form / You behold in one body / Wonder at its beauty'. Translated by author.

[79]"A double heart in one breast / Often you shall find / Beware of treachery".

[80]he moralising poem for Apollo and Daphne famously reads 'Quisquis amans sequitur / Fugitiuae gaudio formae / Fronde manus implet baccas / Seu carpit amaras' ('Whoever, loving, pursues / The joy of fleeing beauty / Fills the hands with leaves / Or seizes bitter berries'). The discussion of this word-image dynamic extends beyond this thesis. See Bolland, 'Desiderio and diletto', 316 and Van Gastel, Il marmo spirante, 185 - 205.

[81]See also Warwick, Bernini: art as theatre, esp. 92 ' 96, and 'Speaking statues' (2004).

Giovanni Francesco Susini, Sleeping Hermaphrodite, 1639. Bronze, Metropolitan Museum of Art, New York.)

Gianlorenzo Bernini, Apollo and Daphne, 1622 ' 1625. Marble, Galleria Borghese, Rome. Installation view. Image retrieved from Wikimedia Commons.

IV Installation

In 1638, the French painter François Perrier published Segmenta nobilium signorum et statuarum, a collection of engravings of famous antique sculptures.[82] In this publication, an engraving of the Borghese Hermaphrodite by Cornelis van Dalen was featured ' one of its first visual reproductions (figure 13). The pillow, pose and drapery of the original statue are carefully followed. This image, however, named 'Hermaphroditus in Hortis Burghesianis', locates its body in a mountainous landscape with hills and trees. Van Dalen's reproduction reflects the variety of spaces in which the Hermaphrodite was experienced in the early 17th-century. This chapter questions how experiences of the Borghese Hermaphrodite's living presence in its Baroque installation were mediated by its spatial contexts.

The Borghese Hermaphrodite would have probably been visited in conjunction with the casino in which it was installed and the villa in which this casino was constructed. Following Camillo Borghese's papal election, Scipione Borghese acquired large patches of land around a vigna near the Porta Pinciana the family already owned.[83] This eventually developed into the Villa Borghese, a respectably sized Baroque villa with gardens, fountains and hunting grounds. 17th-century visitors would enter this villa from the palazzi on the Quirinal through the Porta Pinciana, along the ancient Aurelian walls (figure 14). Inside the entrance, a lex hortorum would await them:

> "I, custodian of the Villa Borghese on the Pincio, proclaim the following:
> Whoever you are, if you are free, do not fear here the fetters of the law.
> Go where you wish, ask what you wish, leave when you wish.
> These delights are provided more for visitors than for the owner.
> As in the Golden Age, when freedom from the cares of time made everything golden,
> The owner refuses to impose iron laws on visitors who linger here.
> May the friend find goodwill here in the

---

[82] Perrier, Segmenta nobilium signorum et statuarum (1638).

[83] See Coffin, Gardens and gardening, 146 -150, Kalveram, Scipione Borghese, 51 ' 55. For the Villa's history, see Campitelli, Villa Borghese (2003).

place of the law,
But if anyone with deceit and intent should transgress the golden laws of hospitality,
Beware lest the custodian of friendship"[84]

The custodian presents the Villa Borghese here as an 'Elysium of delight' where the 'fetters of the law' from the city do not apply.[85] In this heterotopic space, a certain 'freedom from the cares of time' may be enjoyed. On the one hand, this lex hortorum implies that the visitor's Arcadian experience of this villa will communicate them with an ideal, eternal, antique past.[86] Domenico Montelatici's 1700 guide to the Villa indeed names it an 'example that revives to our Time the Magnificence and the splendour of [...] what came from the ancient glorious Rome'.[87] On the other hand, this instruction of delight can possibly be interrupted by surprise and awe. Right to the entrance, hidden between the trees, a pavillon stood in the villa, where banquets could be enjoyed along a marble table underneath a fresco of the Feast of the Gods.[88] Here, Scipione Borghese would, for example, unexpectedly let flower petals fall from the ceiling when hosting guests.[89] Other eyewitness accounts of contemporary garden banquets describe how these feasts would be accompanied by surprise appearances of water, music, and actors, as well as the actual food itself.[90] In an heterotopic space such as the Baroque Villa Borghese, the experience of pleasure was thus charged with the potential of theatrical surprise.

A 17th-century visitor would have to get lost in between the trees of the Villa Borghese before reaching the casino (figure 15). The construction of this casino occurred parallel to the villa and must have been finished by 1628.[91] From wanderings through the adjacent gardens only the façade of the casino would have been visible. In the 17th-century, this façade was covered with antique reliefs, creating an architectural resemblance to Baroque ephemeral structures created during feasts and processions.[92] This bringing of archaeological artefacts to the forefront, rooted in the intellectual admiration for their historical significance, charges these objects with the possibility of an historical sensation, or a theatrical, emotional experience in which this history momentarily disappears before the eyes of its spectators.[93]

A walk up a small stairwell would then permit entrance into the casino's loggia, where heads of Roman emperors and smaller sculpture would be exhibited.[94] This openness ensures continuity from the delights of the garden surroundings. Once inside, visitors would encounter not just sculptures and paintings, but musical instruments, clocks, and automata.[95] On his 1644 visit to the casino, John Evelyn was shown 'a chair that [...] surprises a man on the sudden, locking him in by the arms and thighs'.[96] Each room was named after the one major work of art that was installed there, usually against the wall.[97] In the room dedicated to Bernini's Apollo and Daphne stood a 'frontispiece [...] with shutters in the middle [which] open[s] and out jumps a terrifying monster's head, which screeches with a terrible voice' echoing Daphne's emotional state.[98] The casino of the Villa Borgh-

---

[84] Manilli, Villa Borghese, 159. Translated by David Coffin.

[85] Evelyn, Diary, 115 ' 116.

[86] Warwick, Bernini: art as theatre, 96 - 101.

[87] Montelatici, Villa Borghese, 1.

[88] Coffin, Gardens and gardening, 147 -149.

[89] Ehrlich, Landscape and identity, 40 - 44.

[90] As described by Carlo Fontana in his personal correspondence. See Coffin, Gardens and gardening, 230 ' 231.

[91] Kalveram, Scipione Borghese, 51 ' 55.

[92] Fagiolo dell'Arco, La festa barocca, 82, and Herrman Fiore, 'The exhibition of sculptures', in Penny and Schmidt, eds., Collecting sculpture, 219 ' 245.

[93] Piranesi, in his design for the façade of Santa Maria del Priorato, incorporated hieroglyphs and reliefs in his design to an arguably similar effect. See Van Eck, Art, agency and living presence, 144-157.

[94] For a plan of the casino, see Kalveram, Scipione Borghese, Appendix III. For the oldest known description of the casino's interior arrangement, see Manilli, Villa Borghese (1650).

[95] The casino never featured any apartments or private rooms.

[96] Evelyn, Diary, 116.

[97] An art work's standing against the wall, as is commonly presumed, does not complicate this art work's capacity to engender living presence responses.

[98] Manilli, Villa Borghese, 70, and Evelyn, Diary, 116. Translated by Genevieve Warwick.

[99] For recent insights on sculpture, experience and musealisation, see Van Eck, eds., Idols and museum pieces (2017).

ese was effectively not a modern museum and it should not be held up to that standard.[99] Rather, 17th-century visitors to the Borghese casino would have their expectations constantly challenged by installations that demanded responses to objects as if they were human beings.

From the imposing rooms on the ground floor, a spiral stairwell would bring visitors to the first floor of the casino. At the top of these stairs, they would reach a gallery in which busts of Roman emperors and statues of gods were displayed. On the other side of this gallery, visitors would finally find the room of the Borghese Hermaphrodite. The statue was probably displayed with the side of the phallus against the wall, its buttocks facing the visitors, its feet being immediately visible from the angle of the gallery (figure 16). The ensemble of the Hermaphrodite's body and mattress was displayed on top of a walnut cassone, decorated with floral patterns, putti and the Borghese coat of arms.[100] This was probably covered with a coperchio, to be opened on the occasion of visits to the casino.[101] Opposite to this installation, a mattress probably similar to Bernini's would have stood.[102] Although the installation of the Hermaphrodite marked the centrepiece of this room, heads of Roman emperors were displayed along the other walls, and a pietra dura table with small statues of Venus vulgaris and Venus pudica on each side of the entrance wall. Large windows, which could have been shut, would have allowed the vivifying dance of light around the Hermaphrodite's body. The display of antique sculpture in the direct surroundings to the Borghese Hermaphrodite creates the expectation of seeing an archaeologically significant object. However, the dynamic shapes of the Hermaphrodite's body and the illusionism of the mattress on which it lays stimulates beholders to walk through its exhibition space. As visitors would have been stimulated to move in a room in which light would have moved, they would have experienced an artistic illusion of movement as if it was not an illusion.

In July 1614, the Spanish ambassador to the

Holy See visited the villa as a guest of Scipione Borghese. He briefly recorded his experience in his diary. During his visit to the 'vigna of signor cardinal Borghese outside the Porta Pinciana', he judged its palazzo the 'most beautiful and agreeable creation of Rome'.[103] The entire day, the ambassador and his famiglia spent enjoying the wine grotto and 'some things sent of edible nature' by the cardinal himself.[104] This anecdote reflects how the villa, the casino and the Hermaphrodite would have primarily accessible to friends and political allies to Scipione Borghese.[105] It is imaginable they would have enjoyed a tour, perhaps personally accompanied by Scipione himself, in which walks through the garden and banquets would be combined with a visit to the casino. It is documented, for example, that Lelio Guidiccioni, Borghese court intellectual, conducted tours of the family's art collections to visitors.[106] Such tours would have created experiences of heterotopic delight through the orchestration of theatrical surprises, whether in the garden or in the casino. The ideological agency of cardinal Scipione Borghese thus ensured the homogeneity of experiences between villa, casino and the Borghese Hermaphrodite.

On October 2, 1633, Scipione Borghese died, 12 years after the end of Paul V's papacy. His descendants inherited the honourable family collections, yet lacked the capacity to expand them. The villa and casino may already have been accessible to artists and curiosi in the time of Scipione Borghese.[107] After his death, the accessibility of the villa and casino to travellers on the Grand Tour must have rapidly expanded. The first reproductions of the Borghese Hermaphrodite as well as mentions in travel guides date from these years immediately following his death. Since these travel guides do not mention the Hermaphrodite's coperchio, it is probable that the Baroque installation started to be dismantled following its rendering accessible to tourists. As the agency of Scipione Borghese waned, so did the animate capacity of the Borghese Hermaphrodite in the spaces in which it was

---

[100] Manilli, Villa Borghese, 98 ' 99.

[101] See note 8. The practice of displaying painting behind curtains has equally been documented. See Kalveram, Scipione Borghese, 119 - 122.

[102] See note 8.

[103] Cited in Kalveram, Scipione Borghese, 79.

[104] Ibidem, 79.

[105] Coffin, Gardens and gardening, 147 ' 149, and Ehrlich, Landscape and identity, 40 - 44.

[106] See Ehrlich, Landscape and identity, 40 - 44, and Warwick, Bernini: art as theatre, 101.

[107] Kalveram, Scipione Borghese, 79.

seen.

Francois Perrier's engraving of the Hermaphrodite in the Borghese hortus echoes multiple spatial experiences of this body. This background surely reflects the natural landscape from the Ovidian narrative. Moreover, the 'Hortus Borghesianus' recalls the Arcadian space of the Villa Borghese, where 17th-century visitors must have felt theatrical delight in conjunction to their experience of the Borghese Hermaphrodite. In the Borghese casino, and in the specific room of the Hermaphrodite, artworks were installed to instigate surprise and confusion rather than aesthetic admiration. Sforza Pallavicino reminded of the 'sensible effects' that awaken the memory of a subject's reality when interacting with an object. In the spaces in which this object is felt, this memory is created, to be activated once such 'sensible effects' are felt. The spaces in which the Borghese Hermaphrodite was installed created the memories to be activated on the instant of sensing the Hermaphrodite's effects. As the statue would perform on the stage that is Bernini's mattress, the Borghese villa and casino operated as theatres in which visitors would have experienced antique artifice as Baroque reality. Once the agent of such experiences, Scipione Borghese, disappeared from his stage, the immersive installation of his Borghese Hermaphrodite probably followed suit.

Cornelis van Dalen, 'Hermaphroditus in Hortis Burghesianis', 1638. Engraving. From Perrier, Segmenta nobilium signorum et statuarum, plate 90. Retrieved from Getty Research Institute.

Arnold van Westerhout, Plan of the Villa Borghese, 1670. Engraving. From Faldi, Li giardini di Roma, plate 34. Retrieved from Getty Research Institute.
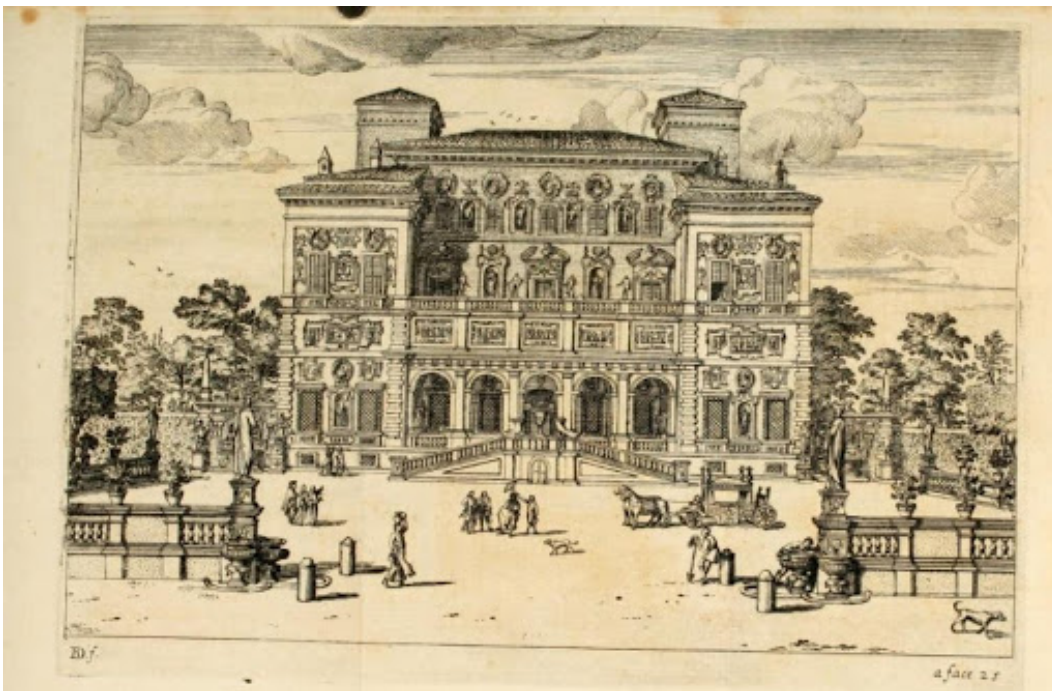
Giacomo Manilli, Façade of the Villa Borghese, 1650. Engraving. From Manilli, Villa Borghese, 25. Retrieved from Getty Research Institute.
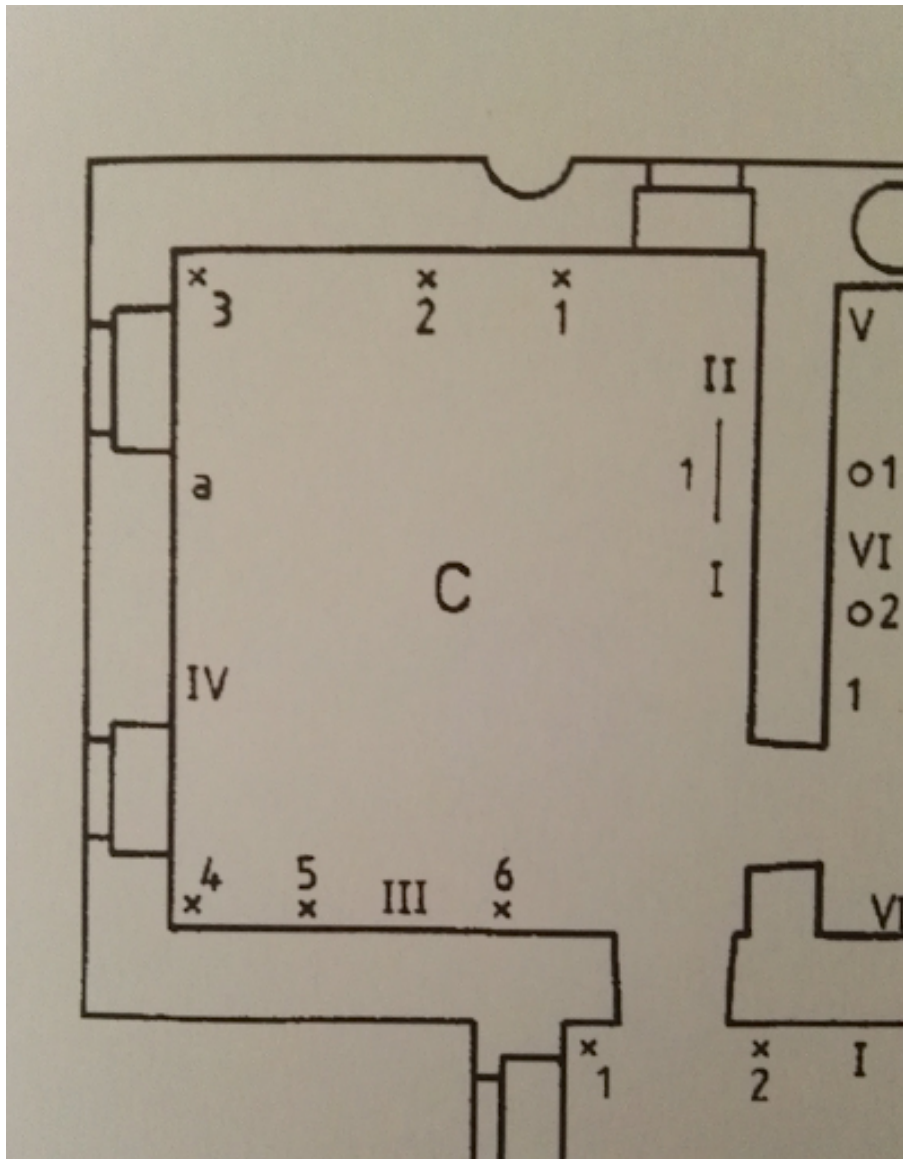
Plan of Room of the Hermaphrodite in the Villa Borghese. From Kalveram, Scipione Borghese, Appendix III. Image courtesy of author. Index: a. Mattress I. Venus vulgaris II. Venus pudica III. Hermaphrodite 1. Pietra dura table 1. Faustina 2. Antonia Augusta 3. Ottacilius 4. Trajan 5. Gallian 6. Consul

V Conclusion

In 1790, Sir John Moore, as one of many Grand Tourists, visited the Hermaphrodite in its 'saloon'. In his published 'view' of 'society and manners in Italy', he decries the Hermaphrodite. Writing that 'the excellence of the execution is disgraced by the vileness of the subject', he expresses his surprise with 'how the Greeks and Romans could take pleasure in such unnatural figures'.[108] In the Borghese Hermaphrodite's life after the Baroque installation, its dynamism seems to have become subject to repression rather than expression. This Capstone has researched how the Baroque installation

of the Borghese Hermaphrodite created an experience of its living presence. In this conclusory section, some central observations from each chapter shall be summarised, before an epilogue meditates on the living presence of the Hermaphrodite's modern life.

The first chapter, 'Discovery', researched how the Hermaphrodite's gender affected the early modern perception of its historical significance. From Ovid's metamorphosis of the nymph Salmacis into Hermaphroditus, a body, neither fully man, nor fully woman emerged that presents an essentially gendered confrontation. The Roman antique

---

[108] Moore, A view, Book II, 376.

sculptural type of the sleeping Hermaphrodite, lifting the buttocks to reveal the penis, follows this location of the Hermaphroditic body in a natural environment. In the early modern reception of the Hermaphrodite, this gendered confrontation was complicated by an historical confrontation. Antonio Beccadelli, in his Hermaphrodite (1425 - 1426), justified his erotic response to its gender by emphasising its historicity. Lorenzo Ghiberti's account equally reflects this tension between his admiration for the Hermaphrodite's history and his desire to touch its body as if it were living. The early modern fascination for the Hermaphrodite's antiquity was already charged with this erotic potential because of its gender.

'Restoration' explored how the restorations executed to the discovered statue of the Hermaphrodite respected its antiquity. Orfeo Boselli proposed the restorer should 'accompany' the antique manner. A fragmented body should be returned to a complete whole. Restorers are most successful, paradoxically, when their artistic imprint is rendered invisible. Besides the nose and hands, the most notable restorations to the Borghese Hermaphrodite occurred around its left foot. The dynamism this body part suggests follows the curves of the Hermaphrodite's entire body, charging it with erotic potential. This aspect specifically contrasts to the Ludovisi Hermaphrodite, a contemporary of the Borghese version, the feet of which lie relaxed on a large rock. This comparison sheds light on two practically divergent approaches to the Baroque practice of restoration. Whereas the Ludovisi Hermaphrodite seems virtually asleep, its Borghese concurrent, as a result of its restorations, is charged with the potential of waking up and becoming alive before the eyes of its spectators.

The third chapter, 'Performance', uncovered how the presence of Bernini's mattress interacts with the historical significance of the statue of the Hermaphrodite. Many experiences of Bernini's mattress include an instant of material confusion. Sforza Pallavicino articulated that artworks can have 'sensible effects' that unconsciously awaken in the mind of the spectator the memory of what this artwork represents. The sense of detail with which this mattress is carved reflects this desire to stage such effects. Bernini's artistic presence thus simultaneously honours and challenges the historical significance of the Hermaphrodite. On

the one hand, his mattress directs the spectator towards the Hermaphrodite, literally and figuratively receding along its weight. On the other hand, this mattress' material confusion charges the Hermaphrodite's fully restored body with the capacity to come alive before its viewers. Bernini thus created what could be considered a stage set to the Hermaphrodite's treacherous performance.

The fourth and final chapter, 'Installation', considered how the spaces in which the Borghese Hermaphrodite was installed mediated this statue's Baroque living presence. The Villa Borghese was developed as an Arcadian space where delight would be enjoyed from the experience of theatrical surprises. In this Villa's casino, the display of automata and musical instruments besides antique sculpture implies an interest in the tension between intellectual and emotional responses. In the room of the Borghese Hermaphrodite, the entry of light could be mediated. The ensemble of statue and mattress was installed with a coperchio, concealing and revealing the Hermaphrodite. 17th-century visitors to the Borghese Hermaphrodite's Baroque installation would have experienced its living presence as a result of their movement in response to the artistic illusion of movement. This theatrical sensation would have probably been activated by a guide on the occasion of a visit by Scipione Borghese's friends or political allies. The agency of Scipione Borghese, patron of the Hermaphrodite, was central to the establishment of this installation. After his death, the theatricality of the Baroque installation gradually eroded.

By 1775, the microcosmic character of the Villa Borghese, its casino, and the installation of its Hermaphrodite had become antiquated. Marcantonio IV Borghese, distant descendant of Scipione, therefore initiated extensive renovations to both villa and casino that would continue until 1800.[109] On this occasion, the Borghese Hermaphrodite was moved to a cabinet on the ground floor, where the statue stood against the wall. As its cassone and coperchio entered disuse, the constraints of this space rendered the Hermaphrodite's penis invisible. Recreated on the occasion of the Hermaphrodite's 2017 return to Rome, this form of display in fact limits the possibility for immersion and interaction around which the Baroque installation was constructed (figure 17).

John Moore's Enlightened repulsions then

---

[109] Paul, Borghese collections, esp. 118 - 137.

emerge as reflections of a shifting cultural understanding of how antique sculpture should be experienced. When in the summer of 1619, a fragmented marble body of an Hermaphrodite emerged from Roman soil, it entered a Baroque forma mentis that actively cultivated the living presence of art works. The Hermaphrodite's patron, Scipione Borghese, and its artist, Bernini, were amongst the main agents of the transformation of this archaeological artefact into a theatrical installation. In the Borghese casino's room of the Hermaphrodite, antiquity came to life, precisely because the experience of artworks as living beings was culturally desired.[110] The rearrangement of the Hermaphrodite in the 1780s then explains itself as the product of Enlightenment aesthetics which value the aesthetic isolation and admiration of art, rather than its vivacity.[111] The imposing architecture of the Louvre's Salle des Caryatides, where the Borghese Hermaphrodite presently resides, follows this same ideological desire. Yet its everlasting pleasant surprises suggest the long-lived Borghese Hermaphrodite may still have a future.

Marilla Sicilla, Borghese Hermaphrodite, 2017. Digital photograph, LaPresse. Installation view. Retrieved from LaPresse.

Works Cited

Primary Sources

Baldinucci, Filippo. La vita del Cavalier Gio. Lorenzo Bernini. Rome: Bernabò, 1713. Retrieved from https://archive.org/details/ gri_000033125010846356.
Beccadelli, Antonio. The Hermaphrodite. Translated by Holt Parker. Cambridge: I

Tatti Renaissance Library, Harvard University Press, 2010.
Brigentius, Andrea. Villa Burghesia, vulgo Pinciana, poeticè descripta, 1716. Google Books. Retrieved from https://books.google.nl/books'id=yP-LCh NyaVUC&dq=andrea+brigentius+villa+bu rghesia&hl=nl&source=gbs_navlinks_s.
Evelyn, John. The diary of John Evelyn. London and New York: Dunne, 1901.
Herder. Johann Gottfried. Plastik: Einige Wahrnehmungen über Form und Gestalt aus Pygmalions bildendem Traume. Riga: Hartknoch, 1778. Retrieved from http://digi.ub.uni-heidelberg.de/diglit/herd er1778/0004/image.
Faldi, Giovanni Battista. Li giardini di Roma: con le loro piante alzate e vedute in prospettiva. Rome: Giacomo de' Rossi, 1670. Retrieved from https: //archive.org/details/gri_33125008634194.
Maffei, Paolo Alessandro. Raccolta di statue antiche e moderne : data in luce sotto i gloriosi auspicj della Santita di N.S. Papa Clemente XI, 1704. Getty Research Institute. Retrieved from https: //archive.org/details/gri_33125012605057.
Manilli, Giacomo. Villa Borghese fuori di Porta Pinciana, 1650. Getty Research Institute. Retrieved from https://archive.org/details/ villaborghesefuo00mani.
Martinelli, Fioravante. Roma ricercato nel suo sito, 1644. Google Books. Retrieved from https://books.google.nl/books'id=wZMEH iP-guQC&printsec=frontcover&hl=nl&sour ce=gbs_ge_summary_r&cad=0#v=onep age&q&f=false.
Montelatici, Domenico. Villa Borghese fuori di Porta Pinciana: con l'ornamenti, che si osservano nel di lei palazzo, e con le figure delle statue più singolare, 1700. Universität Köln. Retrieved from http://arachne.uni-koeln.de/arachne/inde x.php'view%5blayou%5d=buch_item&sea rch%5bconstraints%5d%5bbuch%5d%5b alias%5d=Montelatici1700&search%5bma tch%5d=exact Nibby, Antonio. Monumenti scelti della Villa Borghese, 1832. Google Books. Retrieved from https://archive.org/

---

[110]Further research could explore such installation practices in Roman casini in the Baroque era.
[111]Van Eck, Art, agency and living presence, 119 ' 136.

details/monumentiscelti00borggoog.

Ovid. Metamorphoses. Translated by Frank Miller. Cambridge: Harvard University Press, 1916. Retrieved from Loeb Classical Library.

Perrier, François. Segmenta nobilium signorum e statuaru: quæ temporis dentem inuidium euasere Urbis æ ternæ ruinis erepta, typis æneis ab se commissa perptuæ uenerationis monimentum, 1638. Getty Research Institute. Retrieved from https://archive.org/details/segmentanobilium00perr.

Pliny the Elder, Natural history. Translated by Harris Rackham. Cambridge: Harvard University Press, 1952. Retrieved from Loeb Classical Library.

Raguenet, Francois. Les monumens de Rome, ou, Descriptions des plus beaux ouvrages de peinture, de sculpture, et d'architecture, qui se voyent à Rome, & aux environs, 1701. Getty Research Institute. Retrieved from https://archive.org/details/lesmonumensderom00ragu.

Smollett, Tobias. Travels through France and Italy. Oxford: Oxford University Press, 1919. Retrieved from https://archive.org/details/travelsthroughf01smolgoog.

Visconti, Ennio. Sculture del palazzo della villa Borghese detta Pinciana, 1796. Oxford University. Retrieved from https://archive.org/details/sculturedelpala01lambgoog.

Secondary Sources

Ajootian, Aileen. Natus diformis: hermaphrodites in Greek and Roman art. PhD dissertation, Bryn Mawr College, 1990.

Bacchi, Andrea, and Anna Coliva, eds. Bernini. Milan: Officina, 2017.

Barkan, Leonard. Unearthing the past: archaeology and aesthetics in the making of Renaissance culture. New Haven and London: Yale University Press, 1999.

Beck, Herbert, and Sabine Schulze, eds. Antikenrezeption im Hockbarock. Berlin: Gebr. Mann Verlag, 1989.

Benes, Mirka, and Dianne Harris, eds. Villas and gardens in early modern France and Rome. Cambridge: Cambridge University Press,

2001.

Bober, Phyllis, and Ruth Rubinstein. Renaissance artists and antique sculpture: a handbook of sources. London: Harvey Miller, 1986.

Bolland, Andrea. 'Desiderio and diletto: vision, touch and the poetics of Bernini's Apollo and Daphne'. Art Bulletin 82, no. 2 (2009): 309 ' 330.

Boucher, Bruce. Italian Baroque sculpture. London: Thames and Hudson, 1998.

Brisson, Luc. Sexual ambivalence: androgyny and hermaphroditism in Graeco-Roman antiquity. Translated by Janet Lloyd. Oakland: University of California Press, 2002.

Bruand, Yves. 'La restauration des sculptures antiques du Cardinal Ludovisi (1621 ' 1632)'. Mélanges d'archéologie et d'histoire, 68 (1956): 397 ' 418.

Burnett Grossman, Janet, Jerry Podany, and Marion True, eds. History of restoration of ancient stone sculptures. Papers delivered at a symposium organised by the department of antiquities and antiquities conservation of the J. Paul Getty Museum and held at the Museum 25 ' 27 October 2001. Los Angeles: Getty Publications, 2003.

Butler, Shane, eds. Deep classics: rethinking classical reception. London and New York: Bloomsbury, 2016.

Campitelli, Alberta. Villa Borghese: storia e gestione. Milan: Skira, 2003.

Christian, Kathleen Wren. Empire without end: antiquities collections in Renaissance Rome: c. 1350 ' 1527. New Haven and London: Yale University Press, 2010.

Clark, Kenneth. The nude: a study in ideal form. Princeton: Princeton University Press, 1956. Coffin, David. Gardens and gardening in papal Rome. Princeton: Princeton University Press, 1991.

Coffin, David. The villa in the life of Renaissance Rome. Princeton: Princeton University Press, 1979.

Colantuono, Anthony, and Steven Ostrow, eds. Critical perspectives on Roman Baroque sculpture. University Park: Pennsylvania State University Press, 2014.

Coliva, Anna, and Sebastian Schulze, eds. Bernini scultore. La nascita del barocco in casa Borghese. Rome: Edizioni de Luca, 1998.

Coliva, Anna, eds. I Borghese e l'antico. Milan: Skira, 2011.

Conan, Michel. Baroque garden cultures: emulation, sublimation, subversion. Cambridge: Harvard University Press, 2005.

Delbeke, Maarten. The art of religion: Sforza Pallavicino and art theory in Bernini's Rome. Aldershot: Ashgate, 2012.

Delbeke, Maarten, Evonne Levy, and Steven Ostrow, eds. Bernini's biographies. University Park: Pennsylvania State University Press, 2007.

Dent Weil, Phoebe. 'Contributions towards a history of sculpture techniques: 1. Orfeo Boselli on the restoration of antique sculptures'. Studies in Conservation 12, no. 3 (1967): 81 ' 101.

Di Gaddo, Beata. L'architettura di Villa Borghese: dal giardino private al parco pubblico. Rome: Fratelli Palombi, 1997.

Dreger, Alice Domurat. Hermaphrodites and the medical invention of sex. Cambridge: Harvard University Press, 2000.

Ehrlich, Tracy. Landscape and identity in early modern Rome: villa culture at Frascati in the Borghese era. Cambridge: Cambridge University Press, 2002.

Fagiolo dell'Arco, Marco. La festa barocca. Rome: De Luca, 1997.

Falguières, Patricia. 'La cité fictive. Les collections de cardinaux, à Rome, au XVIe siècle'. Actes du colloque de Rome 106, no. 1 (1988): 215 ' 333.

Feigenbaum, Gail, eds. Display of art in the Roman palace: 1550 ' 1750. Los Angeles: Getty Publications, 2014.

Fragnito, Gigliola. 'Cardinals' courts in sixteenth-century Rome'. Journal of Modern History 65, no. 1 (1993): 26 ' 56.

García Cueto, David. 'On the original meanings of Gian Lorenzo Bernini's Anima beata and Anima dannata: Nymph and satyr" Sculpture Journal 24, no. 1 (2015): 37 ' 53.

Gell, Alfred. Art and agency: an anthropological theory. Oxford: Clarendon Press, 1998.

Gell, Alfred. 'The technology of enchantment and the enchantment of technology'. In The object reader, edited by Fiona Candlin and Raiford Guins. London: Routledge, 2009: 40 ' 63.

Giometti, Cristiano. 'Per accompagnare l'antico. The restoration of ancient sculpture in early eighteenth-century Rome'. Journal of the History of Collections 24, no. 2 (2012): 219 ' 230.

Giuliano, Antonio. La collezione Boncompagni Ludovisi: Algardi, Bernini e la fortuna dell'antico. Venice: Marsilio, 1992.

Gross, Kenneth. The dream of the moving statue. Ithaca: Cornell University Press, 1992.

Hall, James. Dictionary of subjects and symbols in art. Boulder: Westview Press, 1974.

Haskell, Francis. History and its images: art and the interpretation of the past. New Haven and London: Yale University Press, 1995.

Haskell, Francis, and Nicholas Penny. Taste and the antique: the lure of classical sculpture, 1500 ' 1900. New Haven: Yale University Press, 1981.

Henning, Michelle. Museums, media and cultural theory. New York: Open University Press, 2006.

Hollingsworth, Mary, and Carol Richardson, eds. The possessions of a cardinal: politics, piety, and art 1450 ' 1700. University Park: Penn State University Press, 2009.

Kalveram, Katrin. Die Antikensammlung des Kardinals Scipione Borghese. Worms am Rhein: Wernersche Verlagsgesellschaft, 1995.

Lavin, Irving. Bernini and the unity of the visual arts. Oxford: Oxford University Press, 1980.

Levy, Ivonne, and Carolina Mangone, eds. Material Bernini. Burlington: Ashgate, 2016.

Long, Kathleen. Hermaphrodites in Renaissance Europe. Burlington: Ashgate, 2006.

McHam, Sarah Blake. Pliny and the artistic culture of the Italian Renaissance: the legacy of the 'Natural history'. New Haven: Yale University Press, 2013.

Montagu, Jennifer. Roman Baroque sculpture: the industry of art. New Haven: Yale University Press, 1992.

Nagel, Alexander, and Christopher Wood. Anachronic Renaissance. New York: Zone Books, 2010.

Paul, Carola. The Borghese collections and the display of art in the age of the Grand Tour. Aldershot: Ashgate, 2008.

Pierguidi, Stefano. "In materia totale di pitture si

rivolsero al singolar Museo Borghesiano':
La quadreria Borghese tra il palazzo di
Ripetta e la villa Pinciana'. Journal of the
History of Collections 26, no. 2 (2014):
161 ' 170.

Penny, Nicholas, and Elke Schmidt, eds. Collecting
sculpture in early modern Europe. New
Haven and London: Yale University Press,
2008.

Pochat, Götz. Theater und bildende Kunst im
Mittelalter und in der Renaissance in
Italien. Graz: Adeva, 1990.

Reinhardt, Volker. Kardinal Scipione Borghese:
Vermögen, Finanzen und sozialer Aufstieg
eines Papstnepoten. Berlin: De Gruyter,
1984.

Roani Villani, Roberta. 'Precisazioni
sull'Ermafrodito del Louvre e su alcune
sculture appartenuta al Cardinal del
Monte'. Paragone 44 (1993): 37 ' 51.

Rose, Marice, and Alison Poe, eds. Receptions of
antiquity: constructions of gender in early
modern European art: 1300 ' 1600.
Leiden: Brill, 2015.

Saslow, James. Ganymede in the Renaissance:
homosexuality in art and society. New
Haven: Yale University Press, 1986.

Stenhouse, William. 'Visitors, display and
reception in the antiquity collections of
late Renaissance Rome'. Renaissance
Quarterly 58, no. 2 (2005): 397 ' 434.

Summers, David. 'Contrapposto: style and
meaning in Renaissance art'. Art Bulletin
59, no. 3 (1977): 336 ' 361.

Tschudi, Victor. Baroque antiquity: archaeological

imagination in early modern Europe.
Cambridge: Cambridge University Press,
2016.

Van Eck, Caroline. Art, agency, and living
presence: from the animated image to the
excessive object. Boston: De Gruyter,
Leiden: Leiden University Press, 2015.

Van Eck, Caroline. 'Living statues: Alfred Gell's Art
and agency, living presence response and
the sublime'. Art History 33, no. 4 (2010):
642 ' 659.

Van Gastel, Joris. 'Bernini's metamorphosis:
sculpture, poetry and the embodied
beholder'. Word & Image: a Journal of
Verbal/Visual Inquiry 28, no. 2 (2012): 193
' 205.

Van Gastel, Joris. Il marmo spirante: sculpture and
experience in seventeenth-century Rome.
PhD dissertation, University of Leiden,
2011.

Von Stackelberg, Katherine. 'Garden hybrids:
hermaphrodite images in Roman houses'.
Classical Antiquity 33, no. 2 (2014): 305 '
426.

Waddy, Patricia. Seventeenth-century Roman
palaces: use and the art of the plan.
Cambridge: Massachusetts Institute of
Technology Press, 1990.

Warwick, Genevieve. Bernini: art as theatre. New
Haven: Yale University Press, 2012.

Warwick, Genevieve. 'Speaking statues: Bernini's
Apollo and Daphne at the Villa Borghese'.
Art History 27, no. 3 (2004): 353 ' 381.

Wittkower, Rudolf. Bernini: the sculptor of the
Roman Baroque. London: Phaidon, 1955.

The Artificial Truth: Universal Rights in a Post-Truth Culture

Vincent Franz

Abstract

In this paper I explore the potential of social technologies in facilitating or deteriorating the enjoyment and protection of universal rights. In order to illustrate the issue, I will focus my analysis on the recent scandal around the political consulting firm Cambridge Analytica, which appropriated the personal data of millions of Facebook users in order to construct statistical models of personality traits that would be used to target political advertising and to inform political campaigns throughout elections and referenda across different countries. My analysis will draw on a theoretical framework contrasting Habermasian discourse ethics and the Foucauldian notion of discursive power. These perspectives will be connected using works from the field of rhetoric and epistemology in answering the following research question: "Do social technologies and the political systems they engender harbor the possibility for the validation of universal rights?"

Introduction

In science, politics and justice, truth is always the most fiercely contested matter. The status of truth in contemporary culture and its relevance to political affairs is a hotly debated topic. Recently, society has even coined the term post-truth, which is defined as a disregard for facts in favour of personal belief and emotion (Oxford dictionary, web). Post-truth is used to describe a cultural condition of human knowledge, as it posits the general difficulty of discerning truth from falsehood in continuous streams of information. In scholarly discourse, post-truth has been variously connected to the rise of populist movements in western liberal democracies (Speed and Mannion, 2017; Lakoff, 2017; Sismondo, 2017). How and why the term gained traction only recently, and whether it describes a novel phenomenon at all or is simply another perspective from which to perceive the condition of human knowledge, is unclear. In the empirical sciences, truth depends on the interpretation of data which are not infinitely abundant and accessible; this raises doubts if a universal truth, be it a moral or scientific one, can be conceived by the scientific instruments of society and be validated by the discourse it entertains and the political practice it engenders. The idea of such a universal consensus suggests political regimes could form the basis of a universally valid form of justice. The organization of societies around digital information technologies that are globally networked has significantly altered the terrain on which public discourse, the production and exchange of knowledge, operates both in principle and practice. Against a growing abundance of information, however, appears to stand society?s corresponding ignorance of facts. This leads to the emergence of beliefs that defy justification on the basis of facts. This comes to bear on the decisions made by human agents, and consequently the institutions they sustain.

The Universal Declaration of Human Rights (UDHR) claims that all human beings are equally entitled to the same rights and freedoms, in a motion to assure universal peace and brotherhood for future generations. The political and personal rights it sets forth are the standard international set of declarative norms shared by the largest part of the community of nation-states; in this territory the UDHR is, so to speak, the normative and legal rock-bottom of international affairs within the United Nations framework. The articles of the UDHR comprises the personal as well as social rights of the individual that the community of states vow to respect and protect. Article 19 of the UDHR describes the right to freedom of expression and opinion, and links it directly to the freedom to "seek and impart information and ideas through any media and regardless of frontiers". But what are the frontiers of information and ideas?

The content of the declaration aside, in the present moment it is impossible to fully ascertain to what extent the validity of such universal rights is guaranteed by the political systems in place, and if their protection is in fact universally afforded by the institutions that claim to warrant protection of those rights. Today, it is the case that most nation-states have formally subscribed to the declaration,

yet political realities cloud the primacy of the universal declaration. One of the most cynical facts of current human rights regime is that the states adhering to the universal primacy of the UDHR (such as the USA, China, Israel, etc.)are simultaneously amongst the most ardent abusers of these rights.

In this essay, I will investigate whether or not the universal validity of rights is strengthened or weakened by the technologies embedded in discursive practices present in society. The central question behind this analysis will be whether or not the application and use of social technologies incorporate the affordances for equal and inalienable rights within the societies and institutions they sustain. What is the ethical value of these rights, as universal as they may be considered, if they cannot be practically enforced and enjoyed by all? Is the political system negotiated through the discursive practices of social technologies at all compatible with the notion of universal and inalienable rights?

Since the scope of this question exceeds by far the scale of the arguments presented in this essay, I will limit my analysis to the data-practices of the political consulting firm Cambridge Analytica (CA). CA has recently sparked controversy over the extensive acquisition and modelling of vast amounts of personalized social media user data in order to generate statistical models of voter personalities for the purpose of political advertisement. Human rights advocates have already criticized these strategic operations as antagonistic to the principles set forth in the UDHR. In my analysis, I will conjoin two relevant approaches to the services that CA produced for its clients: firstly, I will discuss the ethical dimensions of the practices deployed by CA; secondly, I will return to the premises of the above stated research question and discuss how the use of social technologies affects political discourse and practice. In my approach, I will integrate different theoretical perspectives from the fields of rhetoric, epistemology, moral philosophy, and politics.

After having introduced the main theories, I will discuss their implications for the Westphalian system of sovereign states and their role in the construction and validation of norms. I will use this discussion to proceed into the analysis of the aforementioned case-study. Preceding the conclusion, I will discuss the case-study?s theoretical and practical implication of the universality principle for the UDHR and its institutional superstructure.

Theoretical Framework

All perils of justice revolve around questions of validity. Discourse ethics holds that metaphysical notions of morality are mute in the reality of socio-political arrangements. Proponents of discourse ethics therefore argue that morality and justice are entities produced and validated by political discourse that springs from the autonomous will of individuals. According to Jürgen Habermas, a founding father of discourse ethics, a social contract serves the "institutionalization of the only "innate" right to equal subjective liberties"[1] (Habermas 121, Faktizität und Geltung). In the Theory of Communicative Action (1981) Habermas constructs a theoretical complex that heaves the reasoning over moral rights from metaphysical foundations onto an intersubjective one, defining universal rights and obligations as the necessary consequence of communicative action that establishes their universal validity as the result of discursive reasoning. Habermas argues that rights "define the boundaries within which a subject is entitled to the free confirmation of his will"[2] (Habermas 109, Faktizitaet und Geltung). The "subject" and the "will" this refers to are equivalent to the Kantian conception of the Man of the Enlightenment. In Habermasian discourse ethics, however, this moral outlook cannot be entirely warranted by the subjective will, but rather intersubjective will and pragmatic reason, which require institutionalization to operate and proliferate in order to validate alleged rights and translate them into enforceable norms. These institutions are the product of communicative behaviour that premises independent will; in this view, institutions form the factual manifestation of hypothetical rights.

In democratic discourse, Habermas perceives a fruitful tension between the legality and the legitimacy of legal norms that ultimately renders the rights of autonomous and rational individuals into legal facts requiring enforcement, i.e. flexible norms that govern ethical standards of the social institutions. In the politics of state, these rights are interdependent with laws and the po-

---

[1]In German: "*Institutionalisierung des einzigen angeborenen*" Rechts auf gleiche subjektive Handlungsfreiheiten.

[2]In German: "*Legen die Grenzen fest, innerhalb deren ein Subjekt zur freien Bestaetigung seines Willens berechtigt ist.*"

litical powers that execute them. Furthermore, This tension grants that no norm is set and restricted from scrutiny; rather, it grants the flexibility to apply pragmatic ethical reasoning to potentially arising dilemmas. Habermas' theory of communicative action extensively analyses the relationship between personal rights and the institutional iterations of discursive practices. The principal goal of this is to establish communicative action, the most basic element of human society, as the source of legitimacy for institutions and their normative purposes (Habermas, chapter 3, Theorie Kommunikativen Handelns). To describe the construction of moral consensus, Habermas introduces a distinction between communicative and strategic action, by which he means to segregate the manipulative intent of strategic motives of speaking from the sincere motivations of communicative action to make one?s voice heard and respected in agreeing upon just social arrangements. Habermas identifies three separate validity claims of communicative action: claims to truth (*Wahrheit*), claims to rightness (*Richtigkeit*), and claims to sincerity (*Wahrhaftigkeit*). On the basis of the validity claims to truth, rightness and sincerity, the intersubjective properties of norms are defined. In rights discourse, the separation of these claims enables the potential universal validity of rights claims of the individual onto institutions of justice. This separation constitutes and regulates the norms it universalizes with respect to the expressed dignity of the human person. It entails the idea that institutions are the legitimate product of a social contract that is unambiguously and legitimately subscribed to by the sovereign will of individual subjects affected by the construction of intersubjective norms. The factuality of these norms is embodied in the discourse that surrounds them and the institutions that uphold it. However, the conception of man that purports institutions as extensions of discursive moral sentiments of the public can be problematized as a circular argument that privileges the figure of a fully autonomous, unambiguously rational, Kantian Man (Gaon 717, 1998). The ambitions of Habermasian discourse ethics, therefore, are not without limitations.

Michel Foucault proposes an altogether different perspective on the function of discourse in society: "Discursive practices are not purely and simply modes of manufacture of discourse. They take shape in technical ensembles, in institutions, in behavioural schemes, in types of transmission and dissemination, in pedagogical forms that both impose and maintain them" (Foucault 12, 1997). This assumes that discourse acts to sustain absolute power relations by assigning individuals to a hierarchy of signifiers that imposes order onto society by informing and institutionalizing the normal against the abnormal. In this juxtaposition, institutions and laws are coupled directly with underlying structures of "...a play of prescriptions that govern selection and exclusion" (Foucault 11, 1997). Therefore, the rights protected and enforced by institutions and laws constitute a discourse of exclusion that mirrors power-relationships. Although Foucault and Habermas share the same Marxian premise that socio-political structures historically emerge from the material configuration of the world, they come to radically different conclusions over the moral telos of the discursive formations of institutionalized consensus. Habermas, on the one hand, entertains the notion that the participatory medium of discourse is open to sovereign self-expression that constitutes the factual ground for universal rights to be produced as an intersubjective claim to the validity of those rights, protected and implemented by the force of pragmatic consensus. Whereas this consensus is expressly criticisable by communicatively acting subjects, it also constitutes the factual foundation for the implementation of personal rights as expressive claims to the overarching moral consensus. Foucault, on the other hand, argues the rights expressed in discourse harbour the erroneous notion that discursive rights are true to moral facts he also rejects the separation of claims on truth from claims on rightness. Foucault and Habermas entertain two opposing perspectives of personal rights and their moral functions as they find expression in individual behaviours: the former emphasizes the structure of power-relations that rights are suspended from relativistic power systems between institutions and individuals, the latter by supposing a subjective freedom of will that engenders rights as claims to a social contract and its institutional embodiment. The contradicting assumptions of both views show that ethical norms derived from discourse are prone to conflicting sentiments that challenge the universal validity of claims to truth based on external facts; the moral content of rights discourse is therefore inherently ambiguous and unquantifiable, and thus unknowable.

Moral philosopher Harry Frankfurt identifies a salient aspect of culture that undermines the Habermasian notion of the discourse principle as constitutive of universal moral truth: *Bullshit*. In his treatise On Bullshit (Frankfurt, 1985), he defines as resulting from utterances that are made indifferent of their truth content, which necessarily deprives them from their utility in constructing a pragmatically universal consensus on normative obligations of states, institutions, and individuals. Frankfurt distinguishes Bullshit from lying because a lie implies the liar?s concern for knowing the truth; Bullshit, therefore, is a double-edged sword, as it can misrepresent the state of the world or the speaker?s state of mind simultaneously without a discernible intent to foster falsehood. As a consequence, Bullshit constitutes a discursive fact without explicit truth content. In Frankfurt?s view, Bullshit is a form of misrepresentation enacted with a motivation of pretence. This pretence entails a dual misrepresentation that can hardly be falsified objectively: on the one hand, the speaker of Bullshit misrepresents the world around him, while also misrepresenting his own feelings and attitudes toward this fact, resulting in speech-acts that are nothing but "hot air" with respect to the matters of fact they address. Politicians running for public office commonly have natural inclinations toward speaking Bullshit in order to convey a certain image of themselves, their motivations, and in order to persuade their audiences of their own person, beliefs, and ideology. Frankfurt himself gives only a few anecdotal examples of Bullshit in his own work, to demonstrate Bullshit in speaking situations. When Frankfurt wrote the essay, the internet was still in its infancy, and the gravity it would attain in changing cultural and normative discourse must have been only vaguely conceivable. In the present day, however, it becomes evident that Bullshit is not only a "salient feature" of culture, but potentially an element of its essence. Arguably, Bullshit has become more prolific than it previously was, as the affordances of digital technologies facilitates and encourages the easy dissemination and replication of speech acts without demanding sufficient justification or reflection of its contents. To some extent, the internet affects everybody?s life, and although -or perhaps precisely because - so many use it on a daily basis, it is apparent that it runs rampant with Bullshit.

Although Frankfurt's theory of Bullshit does not offer an explicit account of the validation of rights, it does point toward a root cause of what is now known as post-truth. But what role does Bullshit play in processes of societal formation and, more importantly, the determination of the rights of the person? From the definition offered by Frankfurt, it can be concluded that as a rhetorical device, Bullshit has a controlling function in that it asserts a sense of community and belonging, i.e. individuals produce it to manage appearances and ideologies. Frankfurt admits that Bullshit, despite its vulgar connotation, can indeed be "carefully wrought" (Frankfurt 22) and thus easily disguise itself in convincingly sincere statements and beliefs. This ambiguity present in Bullshit leads him to conclude: "Our natures are indeed elusively insubstantial ? notoriously less stable and inherent than the natures of other things. And insofar this is the case, sincerity itself is Bullshit" (Frankfurt 67). The extent to which sincerity is Bullshit is detrimental to the central function of sincerity in discourse ethics and the process of norm construction it proposes. In order to satisfy claims to sincerity, a speaker must hold true to their word until proven otherwise, making them ?fundamentally criticisable? (Habermas 85, Theorie Kommunikativen Handelns). Following the conceptual caveats of Bullshit, however, a speaker is not necessarily concerned with the truth content of their sincere utterance. Where sincerity is Bullshit, criticism is null and void, because sincere Bullshit can be used to prevent criticism from arising in the first place. Frankfurt?s conclusion thereby challenges the notion that the moral requirements of universal rights are expressed in the factuality of discourse, as sanctioned by the sincerity of the expressed beliefs. The extent to which sincerity is Bullshit consequently undermines the Habermasian trinity of validity claims - the claims to truth, rightness and sincerity - that need to be satisfied before a normative statement can transcend to universal validity. Against this, Habermas could argue that the association between sincerity and Bullshit would be the discernible result of strategic action. However, Borman (2011) develops this argument in reflection of the theory of communicative action, finding that Bullshit "appears to explode the distinction between communicative and strategic action" (Borman 134). This leads him to conclude that the notion of Bullshit probably is Habermas' worst nightmare, because in a society where Bullshit has become sufficiently dominant, such a criti-

cal theory would have become impossible (Borman 135). Where Bullshit dominates the debate, it is impossible to determine whether individual behaviour is in line with normative or moral contemplation of one?s agency. Most sincere utterances cannot be objectively validated; if somebody speaks Bullshit, then the speech-imminent obligations of such a speech act cannot be validated as normative assertion.

Frankfurt's essay was published as a book and gained increased attention from academic and mainstream audiences, and a growing body of literature across various disciplines has been based on its theoretical conceptions; among these applications are both quantitative (Pennycook et al.) and qualitative methods of analysis (Fredal 2011; Wakeham 2017), political philosophy (Borman 2011; Belfiore 2009), and aesthetics (London 2013). The point of convergence between these arguments is that language can and does function independently of truth as such. But is this an impediment to the potential validity of discourse? This leads us back to the idea of a post-truth society as it has emerged more recently: subjective emotions and beliefs, rather than adherence to facts, validating misconceptions and misrepresentations of the world as true. Though the term only surfaced recently, its definition suggests that it is not a novel symptom at all, but a permanent condition that is empirically reinforced by the cultural entertainments of contemporary politics. The growing expanse of information infrastructure pushes the frontiers of the epistemic landscape around the local perception and knowledge of the subject. It becomes clear that the notion of post-truth as it recently emerged is applicable to civilized societies from the first day they recognized their civility.

The theoretical perspectives of Foucault and Frankfurt challenge the Habermasian notion that the universal validity of discursive consensus is the substrate from which truth is engendered. If claims to truth cannot be separated from claims to rightness and sincerity, then the validity of norms, rights, and obligations cannot be expressed and universalized in a discursive formation. The tension between discourse and validity exposes the inadequacy of the notion that truth is grounded in facts in the world that can find universal validation. This limitation of realism can be identified in

Moore's paradox: a fact inferred with a belief that contradicts this fact can be true, logically consistent, and does not necessarily entail a contradiction (Wittgenstein 2009, 190). Habermas cites this paradox to show that claims to truth can be objectively separated from claims to sincerity, holding that validity claims can be verified by independent judicial authority, and sanctioned by discursive justification: "A judge, given sufficient evidence, could in one case criticize the sincere utterance as untrue, and in the other the true expression as insincere"[3] (Habermas 420, Theorie Kommunikativen Handelns). However, Moore?s paradox can be read in reverse, namely that any argument misrepresenting a fact can, regardless of intention and for all practical purposes, be a contradiction: in a culture where Bullshit is salient, uttering it can lead to situations in which the sincere utterance is perceived as true or profound, even if in reality it is false or misleading. The contradictions this produces are also a source of conflict formation (Galtung 1996, 70 ff.).

The objections of Frankfurt and Foucault expose the limitations of the Habermasian notion of truth: Frankfurt shows that speech acts do not always have to serve the validation of a claim to truth about the world or one's state of mind, while Foucault holds that the discursive selection of objects, signs and expressions taken to represent truth cannot be validated as right; Moore?s paradox holds in all cases. This undermines the potential universalizability of validity claims and ultimately challenges Habermas? concept of truth as the product of validation through discursive argumentation relative to a pre-existing reality and its facts. This externalization of truth onto reality at large limits the extent to which it can take root in hearts and minds, and the institutions brought forth from the same. In light of this limitation of universal validity, a new qualification imposes itself when considering the rhetorical nature of the discursive construction of truth and the universal validity of human rights: **artificial truth**.

---

[3]In German: "*Ein Richter, der über hinreichende Evidenz verfügte, könnte in einem Fall die wahrhaftige Äusserung als unwahr kritisieren, im anderen Fall die wahre Äusserung als unwahrhaftig decouvrieren*"

Sovereignty games: Playing with Truth & Playing with Fire

Politics is a game of numbers. In political affairs, numbers yield ?unmistakable political power within technologies of government? (Rose 197, 1999). They determine who holds power, they operate as diagnostic instrument within political reason, and they make government possible and ?judgeable? (ibid.). The same holds true for the information industry and the role it plays in today?s society. The numbers processed by the information industry mobilize populations in the formation of social, economic and political systems that compose the state, and enable institutions to use the power of numbers to implement what Foucault dubs as "the conduct of conduct" (Foucault qtd. in Rose 3, 1999).

Foucault describes historico-political discourse as "a discourse in which truth functions as a weapon" (Foucault 63), challenging Habermas' assumption that the separation of validity claims is the principal foundation for the potential universalization of human rights, as it constitutes the outlines of the autonomous and dignified individual. From the assumption of individual autonomy arises the notion of the state as deriving legitimate sovereignty from the will of its people. In discourse ethics, this constitutes the a priori ground for basing morality in intersubjective claims that constitute the sovereign state, or a federation thereof, as guarantors of an institutional order that facilitates and supports these claims; in the modern state, civil rights are the figureheads of national constitutions, as they map the normative roadmap for basing rights in legal claims to an institutional apparatus. In discourse ethics, institutions are necessary to facilitate a speaking situation in which rights claims can attain validity, and the potential to universalization. Using this line of reasoning Habermas conceptualizes the "ideal speech situation" (*ideale Sprechsituation*) as the hypothetical context for the validity of universal rights in the institutionalization of justice, which constitutes the normative modus vivendi of global moral consensus.

Foucault defends his position in stating that the idea of ?a state of communication that would allow games of truth to circulate freely, without any constraints or coercive effects, seems utopian? (Foucault 1997, 298). Since this argument was made by Foucault, discursive practices have changed radically, and the technological environment through which the subject articulates and expresses itself has significantly changed the ways in which games of truth circulate. Digital technology and the internet revolutionized how humans aggregate and communicate information, and direct it through technocratic institutions sustaining the communicative public. But what do they change about public life? The cybernetic systems surrounding society produce a deep web of information that fragments the experiences of the subject into data that are fed back into the personalized structure of a polymorphous social interface. These data are processed into information that is then utilized in order to garner attention and exhort agency through the same media. The intensified use of information technologies in everyday life significantly changes perceptions of selfhood and identity (Jasanoff et al. 2016) by creating new mechanisms through which publics are formed. In the same way, informational techniques also change how popular sovereignty is informed, exercised, and negotiated in legitimizing politics of the state. Regarding discourses of rights and obligations in this environment Habermas attributes their emergence to liberty (*Handlungsfreiheit*), whereas Foucault attributes it to coercion (*passim*).

Cambridge Analytica

In early 2018 CA appraised its services as the cutting edge way to wage targeted advertising campaigns with greatest efficiency. In its own words, CA describes the impact of its operations on the Trump campaign in 2016 as a "spectacular success" (CA, web), claiming that its scientific approach to digital advertising allowed its customers to influence voters ?when and where it counted?, even if CA "didn't know about their political beliefs" (ibid.). However, the psychographic method used by CA to map the personalities of its target audience is not without limitations: Gonzalez (2017) notes that the controversy around CA is due in large parts to the self-image the company presented to the public, constructing a narrative that exaggerated the success of their services. Nonetheless, CA's strategic employment of psychographic methods, precisely because of its alleged scientific rigour, has to be taken into careful consideration

when assessing the implications of big data technologies for the rights of the individual. Rigour does not imply that the models created by CA are scientifically valuable or significant: the principle purpose of a scientific method can be generally distinguished into epistemic aims, the furtherance of knowledge, and practical aims, the furtherance of utility (Resnik 35, 1998). Upon evaluating CA's method according to those principles, it becomes clear that the "scientific approach" that CA prides itself on is a prime example of Frankfurtian Bullshit. The presupposition that personality traits can be inferred from data produced on social media takes for granted that these data do genuinely reflect personality traits that can be used to identify social and psychological traits of a person; this pretends that digital selves are synonymous with real selves of political subjects exercising their will and political rights. The knowledge thereby generated by CA's model does not satisfy any epistemic aim that is concerned with the production of true knowledge. The practical aim of CA's science is equally dubious; the utility of deriving personality profiles from personal Facebook data implies that open and rational dialogue with the public is superfluous from the outset since subjective beliefs and behaviours are determined by personality traits, excluding the very possibility of there being a (free) will.

Aleksandr Kogan, the Cambridge University professor who programmed the application that harvested Facebook data he exchanged with CA, claimed in an interview that the psychographic method CA applied to his data "gave very little insight in to individuals' personalities", and thus are prone to "large error" in identifying neurotic inclinations? (Sumpter, web). The methods used by CA to derive the ?big five? socio-psychological traits of the OCEAN model from personalized Facebook data, developed by Kosinski et al. (2013), created a statistical confidence around the targeting of political advertisements, thereby influencing the content and context of the speech acts it informed. The knowledge produced by CA, however, does not contribute anything to the quality, or the truth value of the arguments it informs. The personality-profiles generated by CA, so Kosinski claimed, included more knowledge about an individual personality than their spouses could produce (Wade, web). It would be inaccurate to say, however, that the information yielded by these data somehow reveals true insights into the minds of the individuals

they address. The quantified models of personality traits generated from data ?do not inscribe a pre-existing reality, but constitute it? (Rose 198, 1999). This means that data from empirical findings contains facts that constitute reality, which is not equivalent with truth or understanding of the same. Facts, however, do not care about their interpretation: Donna Haraway (1988) insists that all claims to scientific truth are by nature rhetorical and therefore bound to deceive; she calls this the ?privilege of partial perspective?. As boundless as the information processed by CA may have been, each data point only constitutes an abstract quantity representing content interaction; this data is only useful relative to a quantity of other data points. The truth they constitute is therefore rhetorical.

However, the sheer amount of data processed by CA, in combination with the high degree of automation that allowed for the rigorous application of analytical algorithms, has to be recognized, in any case, for the statistical confidence it achieved. Even when assuming the validity of the underlying psychographic model, the strategic intelligence provided by CA's analysis did not contribute anything to the truthfulness or openness of the political discussion. Rather, CA's models identified a specific population that they believed would be the easiest to sway, based on the personality traits they inferred from the acquired data. This does not mean that there is any certainty that the CA operations had no impact on the outcomes of the elections or on the effectiveness of the political or corporate campaigns it informed ? a matter that needs to be investigated at a different place and time. It is important to note, however, that CA, or any targeted advertising for that matter, does not negate the reasoning capacity of the individuals it addresses ? it simply ignores them. The strategic targeting of speech acts does not imply some magic power that sublimely coerces the individual into adopting certain beliefs or behaviours. Rather, it is about finding out what someone would like to read, hear, or see, and subsequently confirming these biases. In a Habermasian sense CA's business practice constitutes strategic action. The speech acts resulting from it, however, are swathed in the proclaimed sincerity of political candidates and their opinions.

Given the amount and scope of the data analysed and utilized by CA in various political campaigns around the globe, their data operations can

be described as a wide-reaching social experiment, as they applied a controlled method to the uncontrolled environment of speech on online social-media. Consent for this experiment, however, was only superficially acquired, which constitutes a breach of most basic ethical research norms; the data of millions of Facebook users were, without their knowledge, acquired by CA and evaluated by algorithms operationalizing the analytical tools at the core of CA's services. This led to the calculation of models that were used to classify and predict behaviour. This information was then used to generate and direct political advertising of its clients in a way that would gain them the greatest support. Following from this, the case can be made that CA intentionally misled voters by suggestively nudging them into a predefined worldview and desired behavioural scheme. Prima facie, this practice seems inhibiting to a constructive and critical function of a truthful political dialogue, that Habermas views as essential to just and legitimate institutions and governance. More severely, it could be claimed that the services provided by CA to political campaigns around the globe have invasively harmed the dignity of the targeted individuals by violating their rights to privacy and depriving their freedom of information and opinion. In a public statement, however, CA claims it acted ?ethically and lawfully?, and named its activities as a ?standard component of online advertising both in the political and social arena? (CA, web). In the latter claim, there is an element of truth, because virtually all internet companies use personal information as their standard commodity because the analysis of such data is at the core of their services. This commodity is valuable because it is a resource of power; as such, it is used to elicit desired behaviours, whether that may be consumer choice, political choice, or the choice of ideology. The advertising force of these assets is valued in the billions. But do these practices engender a genuine concern for ethical issues? CA and its partner corporations were set up and contracted within the sphere of action permitted by civil law and the constitutive norms of the internet infrastructure. Through exploiting the loopholes of Facebook?s infrastructure, CA was able to scrape vast amounts of public and private data from social media platforms without being contested by their unwitting subjects. CA operated through the same infrastructure that drove the elections it supported; by analysing the online livelihoods of individuals, it

used the individuals it affected as means to an end: gaining votes.

Here, it should be reminded that Frankfurt acknowledges that Bullshit is often "carefully wrought" (23). The sophistication or rigour of a method or speaking strategy does not need to have anything to do with the validity of the underlying claims or assumptions. Belfiore (2009) expands this argument by identifying Bullshit as a deep-seated problem in the field of policy research; what she calls "Bullshit of academic variety" is typical of research hypotheses in the field of cultural policy issues that attempt to identify evidence of the attainment of abstract goals. The variety of Bullshit that can be found in CA's method of political consulting is obvious. It entails the assumption that on the basis of comparative social media data, it is possible to determine one?s specific social personality and inclination toward holding certain beliefs that are congruent with ideas or behaviours one expects or wants them to hold, which pretends that these data do in fact resemble intrinsic personality traits. This premise can be compared to the premise that a dating app can help it?s users to find ?love? based on narrow parameters such as looks, location or interest. Likewise, measuring a personality and interpreting the results as a means to predict potential behaviour and beliefs of the respective individuals is equally pretentious ? for the service to be sold, this proposition does not need to be true, it is sufficient to convincingly make it a selling point. The essence of this methodical Bullshit can be identified in PR-business practices aimed at exhorting consumer choice by influencing inferred beliefs and desires. CA appropriated this approach in order to influence not just consumer-choice, but the very choice of political beliefs.

In encounters between individuals and institutions, Bullshit causes tension between the coercive power of discursive practices and the power of the individual will in choice and justification of beliefs. Wakeham (2017) problematizes this as arising from the tension between social knowledge and subjective understanding: "institutional feedback and the resulting meta-understanding thus can create a socially sustained cognitive dissonance"(29). The feedback loop this creates thus compounds confirmation bias with cognitive dissonance. In this loop, it is difficult to determine the exact origin of the Bullshit that is being uttered, preventing the exercise of epistemic vigilance on account of com-

petent testimony. Since "the unsettled nature of epistemic standards is related to the dimensions of social life that prioritize things other than questions of truth or falsity(,) [...] interactional norms and conversational conventions might encourage us to put aside strong epistemic demands" (Wakeham 26 ff., 2017). This analysis shows that Bullshit is not only morally dubious, but also a technical problem in the epistemic setup of society, leading to a polarization of beliefs. The ensuing deterioration of the "forceless force of the superior argument" (Habermas 47, Theorie Kommunikativen Handelns: "zwanglosen Zwang des besseren Arguments") destabilizes the validity of claims to universal rights at its root, namely in the compliance of institutions. Following the revelation of the privacy breaches CA committed by acquiring and processing data of at least 86 million Facebook-users while foregoing full and active consent caused an international scandal fuelled by concerns over privacy and the independence of the democratic process. It also raised concerns over the ethical pitfalls of social media. It is important to note that the case of CA is only the tip of the iceberg of a greater trend toward intelligent and automated analysis in the utilization of big data processing, in order to more effectively direct advertisement, persuasion, and propaganda, a practice that has over recent years become a commonly known feature of digital consumer-relations and persuasion techniques used on the internet. Former CA research director Christopher Wylie, who blew the whistle on CA, labelled the company's practices as "military-style operations" (web). In his questioning by the British Parliament, he stated that in some instances, particularly in Kenya, CA's intelligence was used to target posts featuring violent and threatening content to the Facebook timelines of opposition supporters (ibid.). Such a strategy is somewhat reminiscent of practices that were and are deployed by authoritarian police states in attempts to quench dissent and exhort popular unity. Another instance where CA's practices could have directly eroded the principles of the UDHR is the 2016 election of Rodrigo Duterte as President of the Philippines, whose campaign allegedly employed the services of CA to scrape the profiles of over a million Filipino citizens. Duterte's campaign was built in large parts on the dehumanization of drug addicts by framing them as a burden for Filipino society. The wave of extrajudicial killings that followed Duterte's election pro-

voked international criticisms by governments and human rights advocates charging him with crimes against humanity and genocide (Simangan). As the international criminal court threatened to investigate the police's involvement in the extrajudicial killings, Duterte revoked the Philippines' membership to the Rome Statute, citing it as an infringement on the sovereignty of the Filipino state. In the context of the CA scandal, Duterte's campaign denies collaboration with CA. To what extent is CA be accountable for being complicit in potential crimes against humanity they did not directly commit, if all they did was reinforcing Duterte's derogatory speech?

Beyond the operations of CA itself, the reaction in the aftermath of their revelation is of central interest when asking whether the technologies deployed in society can incorporate the moral requirements for the universal validity of any right. In response to the CA's data scandal several national democratic institutions initiated inquiries scrutinizing their interference as disrupting civil rights and the democratic processes. At the beginning of this decade, Facebook has been hailed by many as revolutionary medium, serving as empowering infrastructure for grass-roots movements pushing for democratic reform of autocratic regimes in the Middle East, which became known as the Arab Spring. In light of the revelation of the CA scandal, however, the democratic influence exerted through social media has been subjected to newly emerging criticism questioning the benefits of data-driven mass mediation. The human rights scholar Sarah Joseph criticizes the business model of Facebook as wholly incompatible with human rights concerns because of the "unprecedented knowledge base and potential power it grants to private actors" (web). Following the CA data scandal, Facebook was also moved into the sights of the representative arbiters of democracy. US congress initiated an inquiry that questioned Mark Zuckerberg, CEO of Facebook, where a large proportion of CA's data came from. Zuckerberg, who before Congress sincerely described his platform as ?idealistic? (web), emphasized the company?s concern for freedom of expression and strengthening political dialogue and participation. That this may be Bullshit can be seen in comparison to statements made by Zuckerberg in the beginnings of Facebook, where he ridiculed the thoughtless compliance of users in his business scheme before investors (The

Telegraph, web). However, the vision Zuckerberg?s claimed "idealism" purports ignores the commercial interest that he and his shareholders have in the operation of Facebook. The idealism portrayed by Zuckerberg seeks to parallel that of Habermas' paradigm of the ideal speech situation, in which all participants are able to fully and freely express themselves. Like any set of parallels these visions will never touch: while Habermas' vision is hypothetical, Zuckerberg's vision is technical.

The principle of discourse, Habermas argues, is morally neutral as it enables the consideration of anyone?s expressed needs and concerns as it legitimizes a system of laws and rights altogether (Habermas 251, Faktizität und Geltung; Apel, 2007). Karl-Otto Apel, a long-time colleague of Habermas and first theorist of discourse ethics partly disagrees with this view; Apel argues that the discourse principle cannot be morally neutral due to the universalization principle of rights and laws that needs to satisfy the categorical imperative as pre-requisite for the universal validation of moral claims (Apel 2007). This entails that the design and utilization, i.e. the technological assemblage of the architectonics of discourse is an ethically relevant practice in determining the discursive configuration of societies. For a global universalization of the rights principle discursive practices must adhere to the categorical imperative. Though algorithms do not possess an intrinsic sense of justice, or a moral will, the designers creating them, and the people employing them, do. It is their context-dependent appropriation of technologies which bears the consequences. This raises the relevant concern whether the architectonics of discourse are indeed appropriate facilities for "globalizing pragmatic reason" (Apel) in order to satisfy the categorical imperative. Or rather, how would the architectonic of the categorical imperative be constituted by technologies and their context dependent uses?

CA's operations can be described as a service intended to inform the power relationship between commercial and political institutions and the online population. The aggregation of user data acquired from social media sources constituted the informational basis on which they attempted to leverage political speech through the same media. This strategy of "selection and exclusion" (Foucault

1997, 18) of personalities, produced by the analysis of Facebook data, enabled CA to claim to know who was most susceptible to what type of speech, where and when. This mode of disseminating political speech, however, excludes concerns for the rightness of affairs of the state from political discourse. The disruption of privacy in order to manipulate the contextual access to public discourse as a means to control voter-behaviour for political ends challenges the Habermasian notion that claims universal rights have the potential to find universal validation in political discourse. It also excludes subjective capacities for reason *per se*, and therefore eliminates concern for the individual contexts in which these technologies are used.

Universal Rights in Discursive Practice

Social mediation has, in a way, become a new world religion. The god this religion venerates is no more a metaphysical being in the heavens beyond; rather, we became the Gods on Earth. Anyone with access to the internet today produces data driven by their navigation on the information-sphere encompassing cyber-space. The users of these spaces are thus at the same time consumers and producers of the technological structures they co-inhabit. Controlling flows of information is a means of consolidating and exercising power over the formation of public discourse. Discursive practices are the groundwork through which a national community, an imagined community as posed by Benedict Anderson (2006), constructs and engenders the sovereign state. The information produced and channelled through social technology is crucial to directing the flow of ideas necessary to forge and maintain the modern state and the social, commercial and military enterprises nested within it. The use of technology in discursive practice at large therefore shapes the structural framework through which these enterprises revolve. The globally networked complex of information technologies and their capitalization through multinational social media corporations thereby constitutes a state within the state, or rather a state beyond the state, which directs and distracts the attention of governmentality"[4] (Foucault 68, 1997). In this light, the notion of

---

[4]Foucault 1997, 68: "[...] the way in which the behavior of a set of individuals became involved, more and more markedly, in the exercise of sovereign power.

preemptive policing acquires a whole new dimension of meaning. Rather than policing against the deviation from a norm, the construction of regulative norms itself is policed by the capabilities of technological intervention.The idea that the world can be made controllable if subjected to sufficient surveillance and appropriative policing pervaded the rise and fall of many an authoritarian regime in history and might eventually pervade that of current political regimes. The institutional shortcomings of the current international universal rights goals and the UN system at facilitating a universal consensus on questions of morality and justice are implicitly criticised by Peter Sloterdijk, who states in his *critique of extremist reason* "[...] that behind the dominating mock-up of liberal pacifism [is] a thorough polemical practice at work."[5] (Sloterdijk chapter 5, Was geschah im 20. Jahrhundert, 2017). Sloterdijk's vagueness in situating this ?mock-up? cannot be blamed solely on his lack of competence in the field of politics, but is rather due to the diffuse sources of Bullshit that pervade the polemic underlying liberal democratic discourse as practised through social institutions and their attached information technologies. In the context of national, international, and multinational institutions, rights discourse as legal discourse thus need to be critically evaluated with respect to the polemical force it exerts onto claims to moral validity and ethical obligations.

The development of human populations into state-societies, national societies, and democratic societies is historically entangled with the emergence of rights discourse, erupting with full force at historical focal points. As such, proclamations of rights are not just political and legal documents but historic ones tied to transitional moments in political attitudes. The UDHR proclaims that all individuals possess an inherent dignity that presupposes agency and the capacity to reason which are the two main elements of Kantian deontology. The rhetorical character of such claims to universal rights is not to be confused with the notion of natural rights, since such rights imply the absence of a normative constitution and regulation. Baynes (1, 2009) therefore advocates the flexibility of a political conception of human rights discourse as "aimed at securing the basic conditions of membership or inclusion in a political society".

Whether the contents of the conditions of membership are fundamentally negotiable is not being sufficiently problematized by this argumentation; it presupposes that the evolving regime of transnational bodies (multinational corporations, treaty associations, NGO?s etc.) is capable of creating secure "conditions for membership and inclusion". The context of these conditions thus transforms the contents of human rights norms and obligations. This view identifies moral rights with the institutionalized enfranchisement of personal rights as legal claims, i.e. the state apparatus and its technical relationship to the social system. Ingram (2009) criticises the discursive understanding of universal rights by arguing that the institutionalization of rights as legal claims threatens to undermine the secure enjoyment of those rights by trapping individuals in abusive structures that deny the secure enjoyment of their rights. He concludes that this can be resolved "if and only if human rights are conceived as moral aspirations and not simply as legal claims"(Ingram 1). In order to support his claim, Ingram uses the business model of sweatshops to exemplify how a legitimate form of employment prevents the secure enjoyment of the human right to subsistence. In the context of data-driven social media, the sweatshop model also applies. The users create the product, while the platform owners reap the profits from marketing the attention of their users to corporate and political actors.

As shown in the previous section of the text, the case of CA specifically and the proliferation of big data practices more generally manipulate the symmetry of the relationships between institutions and individuals; the "speech situations? they create are characterised by their informational asymmetry. Fredal (2011) argues that encounters characterised by such asymmetries are most prone to lead to the utterance and proliferation of Bullshit, which distorts subjective understandings of the world in interacting with institutions. His argument implies that there is an intrinsic connection between the idea of post-truth and the Bullshit Frankfurt perceives as salient to our culture. In a pre-Frankfurtian definition of Bullshit cited by Fredal, the condition assumed as the root cause of Bullshit is named *Eichmannism* (Postman qtd. in Fredal), which signifies the blind loyalty and bureaucratic zeal that made Adolf Eichmann the lo-

---

[5]In German: "*..wie hinter den heute dominierenden Attrappen des liberalen Pazifismus eine durchwegs polemische Praxis am Werk ist.*"

gistician of the Holocaust. The "banality" Hannah Arendt attributes to the evils committed by Eichmann can be found in the role CA has played in the political context; neither Eichmann nor CA was the original cause of the ensuing consequences. As mentioned above, it is difficult to determine the gravity CA's mass transportation and transmutation of information has exerted on later events, although CA itself proudly claimed a prophetic grasp of those events. The motivation behind the algorithmic recollection of millions of digital selves to derive social personality traits as a basis in order to inform campaign rhetoric reiterates the deeds of Eichmann at a level that transcends bodily and territorial frontiers. As such, it rests on a slippery slope.

The extent to which the supranational institutions of the human rights regime currently in place are capable of enforcing universal protection of these rights is limited by the global status of diplomatic and political consensus; claims to national sovereignty are thus the highest hurdle on the path to universal consensus on issues of universal justice. But is the political will emerging from discursive practices in contemporary democratic culture at all capable of finding such a consensus? A core legal concept of international justice is the notion of *hostis generis humani* (foe of all mankind), which defines certain criminal acts that violate the basic principles of any society and are, therefore, prosecutable by any sovereign power (Burgess 2005). This notion was first applied to the outlawing of pirates, whose crimes at sea were seen as a threat to all powers engaged in maritime trade (this norm, however, was easily circumnavigated, as pirates began to acquire licenses and act as contracted freebooters). Eichmann?s overseas arrest through Mossad agents was sanctioned by the argument that his bureaucratic contribution to the Holocaust rendered him a *hostem generem humani*. To what extent, then, can the CA conglomerate or those involved be perceived as hostis generis humani? Their technological intervention in democratic processes across several national elections and referenda is an element that suggests that Big Data practices could be hostile to the democratic principles and political rights upheld by these states. As such, they inhibit the communicative validation of moral consensus through the peddling of plebiscite. The aforementioned analogy to Eichmann suggests that this answer is not unambiguous. The crimes

Eichmann was charged with were committed from an institutional context in which he merely facilitated, perhaps recklessly, the execution of a ?final solution? that did not originate in his own thinking and behaviour but were in compliance with institutional authority that prescribed it. It is unclear, who, in context of the CA scandal fills the role of Eichmann: Zuckerberg, Kogan, Nix, Mercer, to name but a few, all qualify for the role. Likewise, it would be a stretch to charge CA with legal liability for the populist sentiments and false beliefs it catered to, though it certainly acted as substantial agent in the systematic denial of reality that the figureheads of such movements purport in present-day politics.

Conclusion

It has become clear throughout this argument that the communicative faculties of individuals and their institutional iterations are apt at misrepresenting the world as it is. Using Frankfurt?s theory of Bullshit, I have attempted to show that such benign misrepresentations have profound influence on social conceptions of truth and their institutional and political manifestations. In combination with Foucault's notion of discursive power informing governmentality, Bullshit offers an explanation of how truth is weaponized in political discourse as claimed by Foucault. The power of this weapon rests not only in withholding or determining the truth but rather in making concern for the truth irrelevant in discourse practice. In this respect, Frankfurt?s theory of Bullshit offers a sharp critical perspective on the shortcomings of critical theory itself, thereby challenging the Habermasian notion that the moral requirements of universal rights could be validated by discursive practices that rely on an externalization of claims to truth, and their separation from claims to rightness and sincerity. The coercive power Foucault perceives in discursive practices is driven not solely by what is held to be true, but by what is pretended to be true. The difference between the two entails there is a distinction to be drawn between Frankfurt's analytical concept and Foucault?s theoretical deliberations, though with respect to the shortcomings of discourse ethics, their ideas complement one another?s. Both directly challenge the legitimacy and authority of social institutions in enforcing a moral

consensus that could claim universal validity of innate and inalienable rights. Ultimately, the combined criticism of Foucault and Frankfurt challenge the primacy of discourse in the validation norms and rights, as the technologies at the core of discursive practices are no neutral ground for establishing a normative consensus in the first place.

The term post-truth has risen following the collective experience of political events that defy common sense, and it resonates with a stress-reaction to the thoroughgoing proliferation of Bullshit through the technological hallmark of discursive practices. Paradoxically, it seems that the idea of post-truth is itself Bullshit, because its meaning cannot be locally or temporarily identified nor related to a factual, external truth that has authority over opinions; this also unveils the flaws of the premise that the truth can be inferred from factual data. The moment of prominence post-truth achieved changes nothing about the reality humanity co-inhabits. What this moment describes is the banal polemic that pervades the appropriation of discursive practices and the consensus they produce on which the divisions and ideologies thrive through the very institutions meant to unite people in striving for peace and brotherhood. The social technologies at the core of these institutions are not the source of the problem as such, but rather their contextual application in pursuit of this polemic.

The method and scale of big data industries is exemplary of the coercive capabilities of social technologies by means of surveillance and selective targeting of information. Whether or not such technological interventions are at all successful at causing the effects they promise, it is important to consider that the purpose to which they are deployed hampers the unconstrained and transparent participation in political and moral legitimation by eliminating concern to reach a valid consensus on morality and justice in the first place. The discursive practices of Big Data admittedly do have a heavy hand in asserting modes of behaving, communicating, and situating individuals into the digital sphere. The society of digital natives this creates gathers around models of knowledge that impose new constraints with the promise of greater freedom. The systematic selection of information in an attempt to exhort opinions, beliefs, and desires in an automated manner prohibits the very notion that these models of knowledge are con-

ducive to a universal or at least pragmatic consensus on moral rights and obligations that transcends the frontiers of cultural and institutional divisions. On the grander scale of the current development the technological setup of discursive practices is determining modes of participation in public reasoning and the practice of norm validation and implementation. That these practices are antagonistic to the aspiration of a moral consensus of universal significance might work as a violent awakening. Perhaps, we will then realize that the right to truth belongs to all people.

Works Cited

Primary Sources

Universal Declaration of Human Rights, United Nations, New York, web.

"Post-Truth". in Oxford Dictionary online, web.

Wylie in Parliament, Guardian News, YouTube, uploaded 27.04.2018.

Cambridge Analytica Website, CA, web, last retrieved 12.05.2018.

Zuckerberg in Congress, Guardian News, YouTube, uploaded 10.04.2018.

Joseph, Sarah. "Why the Business Model of Social Media Giants like Facebook is Incompatible with Human Rights". The Conversation, web, last retrieved 23.05.2018

Sumpter, David. "My Interview with Alexandr Kogan: What Cambridge Analytica were Trying to do, and Why their Algorithm Did not Work", Medium, web, last retrieved 23.05.2018

Wade, Michael. "Psychographics-the behavioural analysis that helped Cambridge Analytica know voters' minds". Phys.org, web, last retrieved 23.05.2018.

The Telegraph. "Mark Zuckerberg mocked Facebook users in student messages." web, last retrieved 28.05.2018

Secondary Sources

Anderson, Benedict R. OG. "Imagined Communities." Reflections on the Origin and Spread of Nationalism. Verso, 2006.

Apel, Karl Otto. "Discourse Ethics, Democracy, and International Law: Toward a Globalization of Practical Reason." American Journal of Economics and Sociology, vol. 66, no. 1, 2007, pp. 49?70, doi:10.1111/j.1536-7150.2007.00497.x.

Baynes, Kenneth. "Toward a Political Conception of Human Rights." Philosophy and Social Criticism, vol. 35, no. 4, 2009, pp. 371?90, doi:10.1177/0191453708102091.

Belfiore, Eleonora. "On Bullshit in Cultural Policy Practice and Research: Notes from the British Case." International Journal of Cultural Policy, vol. 15, no. 3, Routledge , Aug. 2009, pp. 343?59, doi:10.1080/10286630902806080.

Borman, David A. "Bullshit, Social Integration, and Political Legitimation: Habermasian Reflections." Dialogue-Canadian Philosophical Review, vol. 50, no. 1, 2011, pp. 117?40, doi:10.1017/S001221731100014X.

Burgess, Douglas R. Jr. "Hostis Humani Generi: Piracy, Terrorism and a New International Law." University of Miami International & Comparative Law Review, vol. 13, 2005, https://heinonline.org/HOL/Page?handle=hein.journals/miaicr13&id=307&div=&collection=.

Foucault, Michel. "Ethics: Subjectivity and Truth." Essential Works of Michel Foucault, 1954-1984. Vol. 1. One, Ed. P. Rabinow, Trans. R. Hurley et Al. London: Allen Lane, 1997.

Frankfurt, H. G. "On Bullshit." Raritan, vol. 6, 1985, doi:10.1080/10584600701641920. Fredal, James. "Rhetoric and Bullshit." College English, vol. 73, no. 3, 2011, pp. 243?59, doi: 10.2307/25790474?ref=search-gateway: 48748c9976b1fa16ca057da2ab494925.

Galtung, Johan. "Peace by Peaceful Means: Peace and Conflict, Development and Civilization." International Peace Research Institute, 1996.

Gaon, Stella. "Pluralizing Universal Man: The Legacy of Transcendentalism and Teleology in Habermas's Discourse Ethics." Review of Politics, vol. 60, no. 4, 1998, pp. 685?718, doi:10.1017/S0034670500050853.

González, Roberto J. "Hacking the Citizenry". Anthropology Today, vol. 33, no. 3, 2017, pp. 9?12.

Habermas, Jürgen. "Faktizität Und Geltung. Beiträge Zur Diskurstheorie Des Rechts Und Des Demokratischen Rechtsstaats." 1995, https://philpapers.org/rec/HABFUG. —. Theorie Des Kommunikativen Handelns. Suhrkamp, 1981.

Haraway, Donna. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." Feminist Studies, vol. 14, no. 3, 1988, doi:10.2307/3178066.

Ingram, David. "Of Sweatshops and Subsistence: Habermas on Human Rights." Ethics and Global Politics, vol. 2, no. 3, 2009, pp. 193?217, doi:10.3402/egp.v2i3.2045.

Jasanoff, Sheila. "The Ethics of Invention Technology and the Human Future." Tantor Media, 2016, https://www.worldcat.org/title/ethics-of-invention-technology-and-the-human-future/oclc/961960316&referer=brief_results.

Kosinski, Michal; Stillwell, David; Graepel, Thore. "Private Traits and Attributes Are Predictable from Digital Records of Human Behavior." Proceedings of the National Academy of Sciences of the United States of America, vol. 110, no. 15, National Academy of Sciences, Apr. 2013, pp. 5802?05, doi:10.1073/pnas.1218772110.

Lakoff, Robin T. "The Hollow Man." Journal of Language and Politics, vol. 16, no. 4, Oct. 2017, pp. 595?606, doi:10.1075/jlp.17022.lak.

London, Justin. ?Ephemeral Media, Ephemeral Works, and Sonny Boy Williamson's "Little Village." The Journal of Aesthetics and Art Criticism, vol. 71, 2013, pp. 45?53, doi:10.2307/23597534.

Pennycook, Gordon, et al. "On the Reception and Detection of Pseudo-Profound Bullshit." Judgment and Decision Making, vol. 10, no. 6, 2015, pp. 549?63, http://journal.sjdm.org/15/15923a/jdm15923a.pdf.

Resnik, David. The Ethics of Science: An Introduction. 1998, https://www.worldcat.org/title/the-ethics-of-science-an-introduction/oclc/812969452&referer=brief_results.

Rose, Nikolas. "Powers of Freedom: Framing Political Thought." Cambridge, Cambridge University Press, 1999, doi:10.1017/CBO9781107415324.004.

Simangan, Dahlia. ?Is the Philippine ?War on Drugs? an Act of Genocide?? Journal of Genocide Research, vol. 3528, no. December, 2017, pp. 1?22, doi:10.1080/14623528.2017.1379939.

Sismondo, Sergio. ?Post-Truth?? Social Studies of Science, vol. 47, no. 1, 2017, pp. 3?6, doi:10.1177/0306312717692076.

Sloterdijk, Peter. Was Geschah Im 20. Jahrhundert? Unterwegs Zu Einer Kritik Der Extremistischen Vernunft. 1. Auflage, Suhrkamp Verlag, 2017.

Speed, Ewen, and Russell Mannion. ?The Rise of Post-Truth Populism in Pluralist Liberal Democracies: Challenges for Health Policy.? International Journal of Health Policy and Management, vol. 6, no. 5, Feb. 2017, pp. 249?51, doi:10.15171/ijhpm.2017.19.

Wakeham, Joshua. ?Bullshit as a Problem of Social Epistemology.? Sociological Theory, vol. 35, no. 1, 2017, pp. 15?38, doi:10.1177/0735275117692835.

Wittgenstein, Ludwig. Philosophische Untersuchungen / Philosophical Investigations. Wiley-Blackwell, 2009.

The Ethics of Fairness in Algorithmic Decision Making: an Ethical Analysis of COMPAS Through Berk's Model of Statistical Fairness

Dorian Buijse

Abstract

Algorithmic decision making tools are used in the judicial system as a way to assess one's risk of recidivism. Although such instruments are gaining increasing popularity in both Europe and America, the use of these algorithms has been heavily debated over the past years. The Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), an American risk assessment, has been shown to treat black and white defendants unequally. This research discusses our moral obligations regarding these algorithms through the lens of normative ethics. Firstly, this study analyzes Berk's different statistical fairness metrics and applies them to COMPAS. Secondly, the practical necessity of each fairness metric is discussed, and our moral obligation to satisfy treatment equality is studied through three theories of normative ethics. Secondly, the practical necessity of each fairness metric is discussed and our moral obligation to satisfy treatment equality is studied through three normative ethical theories, namely consequentialism, namely consequentialism, Kantianism and virtue ethics. Lastly, through the use of thought experiments we attempt to extrapolate our findings for risk assessment to the fairness of more general algorithmic assessments. This research shows how looking at the short term effects allows consequentialists to argue in favor of the current use of risk assessment algorithms; however, in the long run the current use of risk assessment algorithms is highly unethical.

Keywords and phrases: fairness; risk assessment; normative ethics; decision-making tool; COMPAS

Introduction

Algorithms are increasingly used as an expert's decision aid in a multitude of fields, including law enforcement, medicine, employment and education (Barocas2016, Berk2012). In the past years, numerous books and papers were written raising awareness of current algorithmic malpractices, making the use of these algorithms highly controversial (O'Neil:2016b, Angwin:2016b, Berk2017, Eubanks:2018). As Cathy O'Neil puts it, "the algorithms encode human prejudice, misunderstanding, and bias into automatic systems that increasingly manage our lives" (O'Neil:2016a).

The algorithms O'Neil (2016b) refers to are based on the unjust profiling of individuals. Profiling can be defined as any automated procedure in which any person related information (personal data) is statistically evaluated to make predictions about specific aspects of individuals such as one's interests, reliability or behaviour (Article 4(4))eu:gdpr. Profiling is not only used for targeted advertising; it is also used by colleges and businesses in their application processes and by insurance companies and government institutions in their risk assessment of individuals (O'Neil:2016b). One such profiling based decision-making al-

gorithm is the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS).

COMPAS is one of the fourth generation decision-making tools used in the American justice system and helps judges decide whether or not to release a defendant who is awaiting trial. The algorithm is based on 137 personal questions relating to family, peer and personal criminal history, economic well-being, education, and personality (NorthpointeInc.2011). A fourth generation decision-making algorithm consists of two parts: the first predicts a defendant's risk of recidivism, and the second predicts their needs in areas such as employment, housing, and substance abuse (Northpointe2015). The former scores surrounding the 'Risk of (Violent) Recidivism' will be thoroughly discussed, as they supposedly signify the likelihood of a defendant to commit a (violent) crime if released, and can sway a judge to either detain or release a defendant. For example, defendants with a score of 7 are on average predicted twice as likely to recidivate as defendants with a score of 3 (Angwin:2016a). A high score can thus serve as a strong argument to detain a defendant who awaits trial.

Algorithms like COMPAS are the newest developments in a larger set of empirical risk

assessment tools. Simpler risk assessment methods have been used as forecasting tools in the American criminal justice system since the start of the 20th century, and have been adopted and adapted by European countries (Berk2012). These algorithms were created in order to get a better grasp of the recidivating group of defendants. With the arrival of computers and machine learning algorithms, this so called "evidence based sentencing" is flourishing in the United States (Starr2014). American risk assessment algorithms are approved and promoted by the American Law Institute in the Model Penal Code. It states that these risk assessment algorithms "[...] offer better predictions of future behavior than [...] the intuitions of criminal-justice professionals such as judges and probation officers." (Ramo2014, p.33).

While Europe has stricter data protection laws, risk assessment algorithms are still currently used (e.g. England's Offender Assessment System, or the Netherlands' "Recidive Inschattings Schalen" (RISc)). The General Data Protection Regulation (GDPR), introduced in May 2018, is a legal framework of data protection laws restricting the collection and processing of personal data of individuals within the European Union. In its current implementation, the GDPR prohibits the private sector, and allows the judicial system, to handle data signifying protected group membership, such as ethnicity or gender (eu:gdpr).

Using data that contains indirect information about one's protected group is likely to cause bias. As expected, (Angwin:2016b) statistically showed COMPAS to to be racially biased, particularly against black people. Although there were no questions specifically regarding race, the algorithm still found patterns closely related to race (such as education level and financial background) in the questionnaire's data (Skeem2016, Starr2014). They found that black defendants were on average more likely to receive a higher risk score than white defendants (Angwin:2016a). This unequal treatment results in skewed predictions, as dangerous white defendants are given a low risk score twice as often as black defendants, while harmless black defendants are labeled high risk twice as often as harmless white defendants. This gives white people an "unfair" advantage in the judicial system.

COMPAS' developer Equivant (formerly Northpointe) and independent researchers Flores et al. complicated the situation by showing the algorithm to satisfy *calibration*, a known statistical measure of *fairness* (Dieterich2016, Flores2016). Consequently, COMPAS seems to be simultaneously fair and unfair. Further research showed this to be the case: Angwin et al. used a different definition of algorithmic fairness than Equivant and Flores et al. (Corbett-Davies2016). This sparked an academic debate about the possible statistical definitions of fairness and how to study the fairness of algorithms (Kleinberg2016, Hardt2016, Corbett-Davies2017, Berk2017, Chouldechova2017). As seen in section 1.2, there is a mutual exclusivity between two metrics of statistical fairness. This means that regardless of the quality of an algorithm, it will never equally handle minority and majority groups.

Currently, risk assessment algorithms are legally viable, although they are statistically likely to have a protected group bias. In order to study the morality of current day decision-making tools, it is important to study the different fairness metrics through normative ethics. Since a singular ethical lens is prone to oversimplification, this research considers three major ethical theories: consequentialism, deontological ethics and virtue ethics.

This research uses COMPAS as a means to understand and discuss the morality of using decision-making tools in our judicial system, by studying the different fairness metrics through the lens of normative ethics. Moreover, the findings will be extrapolated through a multitude of thought experiments aimed to set up some general approximations of our moral obligations when it comes to the use of decision-making algorithms. This research aims to further enhance the societal understanding of the ethical consequences of the current and future use of algorithmic decision-making.

Algorithmic Decision Making

Many important decisions depend on human predictions: doctors predict the effectiveness of treatment; lenders predict pay-off; managers predict future productiveness of their team; judges forecast the risk of releasing defendants; and colleges predict the admissibility of their applicants (Chalfin2016, Athey2011, Kleinberg2017, Berk2017). Until the second half of the 20th century, these predictions were made clinically; predictions were made on the basis of mental models, personal experience and statistical reasoning (Kleinberg2017). Around the 1950s, this began to change, as the fields of psychology and criminology researched the effectiveness of moving from a clinical to an actuarial decision-making process, a process based on previous data and statistics (Ohlin1949, meehl1954, Dawes1989). The actuarial judgment movement was promising in criminology and spread to a multitude of disciplines. Actuarial decision-making research increased in relevance over time, as statistical predictions tend to get better with more data, and data availability has exponentially increased since the 50's.

In the last 70 years, the field of risk assessment has researched and developed new tools. The first improvement of the actuarial tools was combining both static and dynamic factors in order to assess a person's risk (a dynamic risk factor is, for example, drug use). The third generation (3G) risk assessment tools were not only used to predict one's risk, but also assessed one's criminogenic needs (Guy2012). 3G tools were not very effective in providing guidance to the risks and needs, thus a new generation of risk assessment tools were developed. 4G risk assessment tools were built on 3G tools, and improved the risk and need assessment based on a set of responsivity factors. These responsivity factors, including age, gender, and mental health, play a large part in the prediction and management of one's needs. COMPAS is a 4G risk assessment tool that consists of two parts, the first predicts a defendant's risk of recidivism and the second predicts their needs in areas such as employment, housing, and substance abuse (Northpointe2015).

Decision-making tools in this day and age are employed through the use of algorithms. A computer algorithm is a set of computer instructions that solve a problem or complete a task. Decision-making algorithms are based on machine learning. Machine learning algorithms study a large data set and attempt to discover patterns within the data. These patterns are combined in a mathematical framework, a classifier, which can be used to recognize, analyze and classify new data (James:2014). Machine learning algorithms can label data they have never encountered before, which is why it the field is referred to as artificial intelligence (AI). A thorough explanation of machine learning algorithms, the different classifiers and other related concepts is beyond the scope of this research, but can be found in many introductory books (Berk2012, James:2014). What one needs to know for this research is that algorithmic decision-making is done by placing a new case in a previously created framework and by labeling the case based on the similarity with previous cases.

A clear example of the value of machine learning algorithms can be found in the medical field. In (2017), researchers Esteva et. al trained a classifier (convolutional neural network) with 129.450 clinical images of different forms of Melanoma, a form of skin cancer. They found their neural network to perform on par with 21 certified dermatologists in classifying different types of skin cancers. Esteva et. al mention the scalability and speed of their algorithm which, if distributed to mobile devices, could help reduce the privilege gap in health care.

In the example above, one can see a clear application of machine learning to improve the decision-making process of individuals. It is important to note, however, that even if the algorithm predicts a certain type of cancer, this may still be a false prediction. While the prediction is useful most of the time, further medical examination needs to follow before one can start treatment. Machine learning algorithms are able to make decisions; however, in important decision-making it is often advisable to use its output as an expert's opinion rather than a fact (YangMin2010).

The growth of machine learning algorithms is mostly limited by data availability. Esteva et. al (2017) needed 130.000 tagged images in order to train their classifier to be on par with certified dermatologists; if they had less data, it would most likely have reduced the accuracy of their algorithm. To study the predicted prevalence of machine learning algorithms in our future society, the availability of data seems to be a good starting point. In 2025, the global data sphere is predicted to quintuple, from 35 zettabytes ($35 * 10^{12}$ Gigabytes) of data today to 163 zettabytes (Reinsel2017). This increase in data is promising for machine learning; however, not everyone is happy with the increase in algorithmic power. In the following chapter, the complications of machine learning algorithms will be contrasted with the aims of our current society.

Complications of Decision-Making tools

Machine learning algorithms already heavily influence our society, and hand in hand with the societal increase of data is the prevalence of machine learning algorithms. While machine learning is accepted as a necessary and useful tool by most, some are more cautious about the transition from clinical to algorithmic decision-making. It is important to note the concerns surrounding machine learning algorithms as it is easy to get absorbed by their increased popularity and flow of positives. The main concern is the decrease in transparency of decisions; in complex and opaque algorithms finding the reason for a classification is almost impossible.

Opacity

The justice department has been busy creating algorithmic tools based on machine learning principles (Berk2012, Hardt2016, Kleinberg2017). It is unclear which classifiers risk assessment algorithms use, as the algorithms are not available to the public. This secrecy is necessary because a public algorithm would in theory be 'gameable', because the

algorithm is consistent and thus one can learn to understand what answers result in a low risk score. Studying the behaviour of machine learning algorithms can prove difficult as the classifiers are mathematical and (mostly) unexplainable in words; however, in an attempt to game the algorithm prior to conviction, one could repeatedly change their input data and record the output score in order to understand what answers result in a low score. Even when someone optimizes their answers, they will most likely never be able to understand why the algorithm gave them a specific score.

This is because machine learning algorithms are inherently opaque. An algorithm is transparent when one can easily explain the reasoning behind an output and opaque when one can not. Opacity can make it difficult for one to abuse the algorithm, but it also raises problems in and of itself. Machine learning algorithms are sometimes so complex that it is impossible to give arguments as to why a certain prediction was made. This 'black box' phenomenon makes it, for example, impossible for a judge to understand why a risk assessment algorithm gives defendants a certain risk score (Bunge1963). The effect of black boxing is best described using Latour's words:

> "[Black boxing is] the way scientific and technical work is made invisible by its own success. When a machine runs efficiently, when a matter of fact is settled, one need focus only on its inputs and outputs and not on its internal complexity. Thus, paradoxically, the more science and technology succeed, the more opaque and obscure they become." (Latour1999)

This raises the general question: should we allow the use of algorithms in decision-making in order to increase efficacy if it means increasing the opacity and obscurity of a given system (Kleinberg2017)?

In many areas increasing opacity is not very problematic: when a dermatology algorithm correctly labels a cancerous tumor, understanding its reasoning seems of little importance. To rephrase Latour's words, if an al-

gorithm works better than any other tool we have, but we do not understand why, we will generally allow an increase of opacity in the name of an increase of efficacy.

In the case of risk assessment algorithms in the justice system it is slightly different. An algorithm in the justice system is either available to the public or it is not. If it is not available, the reasoning behind classifications is unexplainable, and the transparency of the system decreases. If an algorithm is available, it is most likely based on complex classifiers such as deep neural networks, which are nearly impossible to translate to human language. Either way, the algorithm's reasoning remains unknown. Therefore, unless the algorithm is based on a naturally transparent classifier, such as a decision tree, and the algorithm is made public, it seems that increasing opacity is unavoidable. This increase in the opacity of our judicial system directly opposes the current Western tendency to strive towards transparent justice (Jaconelli2002).

However, it is important to keep in mind that at some point, a change in judicial structure could occur. It may be interesting to consider foregoing transparent justice if algorithms can increase the efficacy of the justice system and subsequently increase public safety. However, more interdisciplinary research is required in the fields of ethics, law and machine learning before such claims can be made. As will be discussed in the next section, the efficacy of risk assessment algorithms is currently not at a stage where a debate is required. The reality of the matter is that currently these decision-making algorithms are approved in the American Justice system, and thus it is important to study and improve them (Ramo2014).

not publicly available, Angwin (2016a) used the Broward county data and COMPAS' risk scores to make predictions about the inner workings of this black box. It has been shown that COMPAS' overall accuracy is around 65%, a bit better than a coin toss (Angwin:2016b). While it is unclear which classifier is used by Equivant, multiple researchers trained their own classifier on the Broward County data, and found similar accuracies with both logistic regressions and Support vector machine classifiers (Dressel2018, YangMin2010) . While COMPAS requires a 137 item questionnaire, (Dressel2018) trained a logistic regression classifier based on two features (age and number of prior convictions), whose efficacity is on par with that of COMPAS. Additionally, the predictions of COMPAS seems to be no more accurate than those of the general public (Dressel2018).

Both European and the American judicial systems currently approve the use of algorithmic tools in their recidivism assessment. Since there is no current motion for change, it would be naive to believe that the system will simply stop their use. While COMPAS may fail to demonstrate acceptable efficacy; it is important to note that there is a big variability in the accuracy of different risk assessment algorithms (Geraghty2015). Additionally, it is important to keep in mind that COMPAS does more than predict risk of recidivism; it also predicts the needs of the defendants. Although COMPAS's needs assessment is not present in this research and its efficacy is not known, it should not be forgotten while thinking of the tool. Therefore, it is still valuable to study COMPAS as a particular but representative example of current day risk assessment algorithms.

Efficacy

Since Angwin (2016a) published Broward County COMPAS data, risk assessment algorithms have received a lot of attention from the media and the academic world (Dieterich2016,Corbett-Davies2016,Kleinberg2016,Berk2017, Corbett-Davies2017, Chouldechova2017, Dressel2018). Because the COMPAS algorithm is

Fairness

The last complication that will be discussed is the fairness of an algorithm. While fairness is generally a social and ethical concept, the field of statistics mostly uses the term to highlight some technical equalities. It is for this reason that even if an algorithm satisfies some statistical fairness criteria, it may still be ethically unfair; this concept will be discussed fur-

ther in the following chapters.

Fairness, a normative principle, is used to suggest outcomes or actions that should occur; if an action is unfair, it is wrong. Socially, fairness often overlaps with other normative principles such as justice, equity, and equality, and is used by individuals to label actions agreeing with their personal moral rightness. Numerous ethical theories of fairness attempt to define fairness as a subset of moral rightness in order to reduce the overlap with justice, equity, and equality.

This paper will follow some statistical definitions of fairness and later compare them to ethical and social concepts. The different mathematical definitions of fairness are based on equalities in the in and output data combined with some societal rules and laws surrounding protected group membership. There are six commonly used statistical fairness metrics, which will be described in the following chapter.

In the past years, numerous articles have appeared studying the behaviour of statistical fairness (Angwin:2016a, Kleinberg2016, Barocas2016, Flores2016, Hardt2016, Dieterich2016, Berk2017, Chouldechova2017, Corbett-Davies2017,Dressel2018). The different fairness metrics led to confusion in the studies on COMPAS. In 2016, COMPAS was found to be *unfair*, as it suffered a racial bias against black people (Angwin:2016b). Black defendants who did not recidivate over a two-year period were misclassified as 'high risk' nearly 45 percent of the time, while white defendants were only misclassified 23 percent of the time. Additionally, white defendants that did recidivate over a two year period were given a 'low risk' score 48 percent of the time, while black defendants in the same situation were misclassified 28 percent of the time. Angwin (2016b) claimed that COMPAS was *unfair* because false positive and false negative rates differed strongly by race, as can be seen seen in chapter 4 they found that COMPAS lacked *(treatment equality)*. COMPAS' developer Equivant (formerly Northpointe) and independent researchers Flores et al. quickly followed with research that proved COMPAS to uphold *predictive parity*, a widely accepted form of fairness (Dieterich2016, Flores2016).

In other words: they showed that of the defendants with a given recidivism score, black and white defendants had a nearly identical rate of reoffence. Since ethical fairness is binary, the partly satisfying of statistical fairness raises questions about the impact of each metric on our perception of the ethical fairness of an algorithm.

To summarize, while risk assessment algorithms are widely used in the western judicial systems they are not without problems. We saw that: decisions made based on algorithmic predictions are unexplainable due to the opaque nature of machine learning algorithms; the current efficacy of risk assessment algorithms is currently not high enough to carry any weight in court; and the effect of the statistical metrics on the ethical fairness of an algorithm is not well studied. The next chapter will be an in-depth look at Berk's (2017) framework of statistical fairness, the mutual exclusivity of predictive parity and treatment equality, and a study of COMPAS through Berk's fairness metrics.

1 Statistical Fairness

A finding like that of Angwin (2016b) makes people question the current and future use of risk assessment algorithms. In order to mitigate possible special group biases like that of COMPAS, researchers have put forward multiple frameworks to test statistical fairness in the fields of computer science, criminology and statistics (Kleinberg2016, Hardt2016, Corbett-Davies2017, Berk2017, Chouldechova2017). All of the frameworks mathematically say similar things; however, there are clear differences in styles and defined concepts. Since the frameworks mostly differ stylistically, this paper will follow Berk et al.'s (2017) framework of fairness, as it has a good balance between its conciseness and its understandability for statistical laymen.

Before the definitions are given, it is important to state some key ideas so that the definitions can be made more formal. In the rest of this thesis, the different predictions will be classified according to a binary system. A correct prediction of a positive or negative out-

come will be called a true positive (TP) and true negative (TN), respectively. Similarly, a false prediction of a positive or negative outcome will be called a false positive (FP) and a false negative (FN) respectively. Note that a failure is seen as a positive outcome and a success as a negative outcome. The algorithm COMPAS is developed to predict recidivism, thus, predicting recidivism is seen as a positive thing.

Berk's model is based on the idea of a confusion table. A confusion table is build based on complete information namely the binary prediction and the true outcome. A confusion table is used to study an algorithm's performance calculating for example the false positive rate (FPR) or the false negative rate (FNR), to study the algorithms misclassifications of each actual outcome.

	Predicted Failure	Predicted Succes	Conditional Procedure Error
Failure - Positive	TP	FN	FNR $= \frac{FN}{(TP+FN)}$
Succes - Negative	FP	TN	FPR $= \frac{FP}{(FP+TN)}$

Table 1: Berk et al. (2017) uses an alteration of Table 1.

When a confusion table is created to test an algorithm, there are some facts about the used data that heavily influences the . The sample size of a test is given by $N = TP + FP + TN + FN$ and the *Base Rate/Prevalence* of the to be predicted action is given by $\frac{TP+FN}{N}$. These

1.1 Definitions of the Fairness Metrics

First lets define $f(L, S)$ as the true function that links legitimate predictors *L* (e.g. prior convictions) and the protected predictors *S* (e.g. race or gender) to the expected outcome of interest *Y*. Mathematically, this would be given by the expected value of Y given L and S, or more formally: $\mathbb{E}(Y|L, S)$. In the case of COMPAS, this outcome of interest could be seen as arresting someone on probation.

The algorithm's estimation of $f(L, S)$ is mathematically defined as $\hat{f}(L, S)$. Note that $\hat{f}(L, S)$ is generally not a perfect predictor. In the following definitions Berk et al. (2017) proceeds as if $\hat{f}(L, S)$ has no estimation bias and thus perfectly estimates $\mathbb{E}(Y|L, S)$. Using this we can define fairness as follows:

1. *Overall accuracy equality* is attained by $\hat{f}(L, S)$ iff:
   The overall accuracy $\frac{TP+TN}{N}$ is the same for all members of *S*.

2. *Statistical Parity* is attained by $\hat{f}(L, S)$ iff:
   The distributions of the predicted classes

are the same for each member of *S*. More formally, in the binary case of a positive and negative class the values of $\frac{TP+FP}{N}$ and $\frac{TN+FN}{N}$ need to be the same for all members of *S*. **Note:** the two can be different from each other.

3. *Conditional Procedure Accuracy Equality* is attained by $\hat{f}(L, S)$ iff:
   The true positive rate (TPR) $\frac{TP}{TP+FN}$ and the true negative rate (TNR) $\frac{TN}{TN+FP}$ are the same for all members of *S*. Like in *statistical parity* the TPR and TNR do not necessarily need to be the same.

4. *Predictive Parity* is attained by $\hat{f}(L, S)$ iff:
   The proportion of correctly predicted positive outcomes is the same across all the protected groups. That is, $\frac{TP}{TP+FP}$ are the same for all members of *S*. **Note:** *predictive parity* is sometimes called *calibration*.

5. *Treatment Equality* is attained by $\hat{f}(L, S)$ iff:
   The ratio of false positives and false negatives (thus: $\frac{FP}{FN}$ or $\frac{FN}{FP}$) is the same for all members of *S*.
   Note: Another form of treatment equality is through the False Positive Rate (FPR) and the False Negative Rate (FNR) for all members of *S*. (defined by $\frac{FP}{TN+FP}$ & $\frac{FN}{TP+FN}$ resp.). The latter form of treatment equality is more frequently used

(Chouldecova 2016; Angwin et al. 2016; Berk et al. 2017).

6. *Total Fairness* is attained by $\hat{f}(L, S)$ iff: All aforementioned definitions of fairness are achieved.

In order to later look at the ethical implications of the different fairness metrics on COMPAS, it is important to clearly place the different metrics in context. *Overall accuracy equality* states that the prediction accuracy of the population should be independent of protected class membership. Thus, in the binary case of $S \in \{black, white\}$ the overall accuracy of the two groups should be equal. It is noteworthy however that *true positives* and *true negatives* are treated as equally important in the valuation of overall accuracy equality; giving a high risk score to a recidivating defendant is just as important as giving a low risk score to non-recidivating defendants. In general ethical arguments can be constructed vouching for a less equal but more desirable treatment of positive and negative classes.

Statistical parity is the property that requires the proportion of positive classifications of a certain protected group to be equal to the proportion of the total population. For COMPAS, this means that the proportion of black and white defendants predicted to succeed on parole is identical. This fairness does not discriminate between true or false predictions, but showcases the algorithm's group bias or lack thereof. Statistical parity is a form of group fairness rather than individual fairness and has been coined as a cause of individual unfairness in algorithms (Dwork2011). If the base rates differ between protected groups, it seems statistically unfair to treat both groups as equal. On the other hand, making a decision (partly) based on protected group membership could also be argued as unfair for the individual.

Conditional procedure accuracy equality requires the correct classification rate to be equal for each protected class. In the case of COMPAS, it requires that one is as likely to be classified properly no matter what protected class one belongs to. This means that the rate of recidivists getting high scores is independent of race. Conditional procedure

accuracy is a common concern within the field of risk assessment algorithms. The fact that most criminal justice applications fail to satisfy this equality means that there is unequal odds within the American justice system (Hardt2016). Namely, black defendants are classified as being inherently more likely to recidive than white defendants. An important aspect to note is that the prevalence of recidivism is higher in black communities, which has been explained as the root cause for this unequal treatment (Chouldechova2017, O'Neil:2016b, Corbett-Davies2017). It must also be stressed that this difference in prevalence has many causes tied to other inequalities within the society, such as a different rates of poverty, police patrolling and neighbourhood crime.

Predictive parity or *calibration* require a classification to be equally useful independent of protected group membership; conditional on the prediction, is the probability of that prediction equal across protected groups. The importance of calibration lies at the core of the discrete classifications. If the likelihood of recidivism among white high-risk offenders is higher than that of black high-risk offenders, one would need to treat the classification of high-risk differently for each race group. This violates the concept of classification of the entire population through the use of one algorithm in that each discrete output would need to also be tagged with one's race.

Treatment equality requires the ratio of false predictions to be equal, which can be seen as equalizing the cost of predictions. When this ratio is out of balance, a false positive may carry more weight for one protected class over the other. This ratio can also be used to show an algorithms racial bias as it shows whether a group is predicted to be more risky.

Remember that these statistical definitions are all calculable metrics of fairness, and do not directly discuss (ethical) fairness itself, they are merely equalities to stress the algorithm's handling of different protected groups. As seen later, while the metrics do not directly reflect the ethical concept of fairness, they can be used to argue about an algorithm's fairness. Each of the fairness metrics hold

true even in situations when there are more than two outcome classes (Berk2017). For the scope of this research it would however be an unnecessary overcomplication to go more in depth.

1.2 The Impossibility Theorem

Chouldechova 2017 shows that the *base rate*, the *positive predictive value* and the *false positive* and *false negative rates* are related through the following equation.

$$\text{FPR} = \frac{\text{B}}{1-\text{B}} \frac{1-\text{PPV}}{\text{PPV}}(1-\text{FNR}) \qquad (1)$$

In case an algorithm satisfies *Predictive Parity*, but the *Base Rates* differs across protected groups, expression 1 states that it is nearly impossible for an algorithm to achieve equal FPRs and FNRs for differing groups. This implies that the rate of a false prediction can never be equal across protected groups unless the model is perfectly accurate model (FPR and FNR are either 0.0 or 1.0).

The simplicity of expression 1 may not immediately set off alarm bells, however its implications are massive. It shows that the statistical definition of total fairness is unachievable for an algorithm if the *Base Rate* between groups in a *non-perfect* model differ. In practice it is highly unlikely for $\hat{f}(L, S)$ to be a perfect predictor or for the base rates of groups to equal (Berk et al. 2017). Expression 1 thus shows us that unless the actual data changes (*Base Rate*) complete equality is mathemati-

cally impossible. Note, however, that this does not immediately imply that the algorithm is ethically unfair; it merely states that protected groups cannot be treated equally unless the base rate is the same.

Implications for COMPAS

The following chapter will use COMPAS data gathered by Angwin (2016a) to analyze the impact of different algorithmic fairness metrics on Caucasian (white) and African American (black) defendants. In order to do this, the aforementioned confusion tables need to be restructured accordingly so that it reflects COMPAS' results properly. The outcome will be binary classified where $Y \in \{0, 1\}$, with $Y = 0$ to indicate the outcome that a given defendant does not recidivate. To avoid complex situations the ten scores will be categorized as high risk and low risk. Scores will be classified as *High Risk* if $S(X) \geq S_{HR}$ where $S(X)$ is defined as the score of individual X. For simplicity's sake, assume X to be a vector constructed with the 137 items from the COMPAS questionnaire and does not include race.

Using the data gathered by Angwin et al. (2016a) the following can construct table 2.

Note that the base rate in table 2 is different between race groups. The base rate for black defendants is 0.51 and 0.39 for white defendants. Using theorem 1, we expect COMPAS to fail either *predictive parity* or *treatment equality*. The following table shows the calculations for all of the statistical fairnesses.

	Black Defendants		White Defendants	
	Low Risk	High Risk	Low Risk	High Risk
$Y = 0$	990	805	1139	349
$Y = 1$	532	1369	461	505

Table 2: Confusion table of the Broward County's data of All Defendants, Black Defendants and White defendants. Here $S_{HR} = 4$. Adapted from (Angwin et al. 2016b)

Probability	Black defendants	White defendants	95% CI
Overall accuracy	0.638	0.670	[-0.0559, -0.0075]*
Statistical parity	0.514	0.394	[0.0955, 0.1459]*
True positive rate	0.720	0.523	[0.1729, 0.2219]*
True negative rate	0.552	0.765	[-0.2371, -0.1907]*
Posivitve predictive value	0.630	0.591	[0.0135, 0.0633]*
Proportion: FP/FN	1.513	0.757	$\star$

Table 3: Using the large data-set from Table 2 (N = 6150), we estimate probabilities through frequencies of COMPAS and the 95% confidence interval of the estimation of $P_B - P_W$.
∗: a rejection of the null hypothesis $P_B - P_W = 0$.
$\star$: The difference in treatment of black and white defendants is so large that the inequality trivial.

Using Table 3 we estimate COMPAS' fairness metrics for $S_{HR} = 4$:

1. *Overall accuracy equality* is not formally satisfied. Note however, that practically P_B and P_W are close to equal.

2. *Statistical Parity* is not formally satisfied as $0.0955 \leq P_B - P_W \leq 0.1459$.

3. *Conditional procedure accuracy equality* is not satisfied.
As $P_b - P_W \gg 0$ for the TPRs and $P_b - P_W \ll 0$ for the TNR.

4. *Predictive parity* is not formally satisfied as $0.0135 \leq P_B - P_W \leq 0.0633$. Just like with Overall Accuracy Equality one can make the argument that the likely difference in PPV is small enough to be practically equal.

5. *Treatment equality* is not satisfied

6. *Total fairness* is not satisfied.

It is important to note that these estimated fairness metrics are made at the cut off point $S_{HR} = 4$. While 4 seems to still be a low score, Chouldechova (2016) shows that for $S_{HR} \geq 4$ the estimations remain consistent. It is however highly probable that since the differences between the ratios slightly changes, the 95% confidence intervals may too. This could mean that some fairnesses that have a CI close to 0 may still be satisfied.

To conclude, while statistically none of the equalities fully hold at $S_{HR} = 4$, we see a clear distinction between orders of difference, whereas *treatment equality* and *conditional procedure accuracy equality* have differences larger than 20%, and *overall accuracy equality* and *predictive parity* seem closer to holding. We expected COMPAS to fail either *predictive parity* or *treatment equality* by theorem 1 and can see that Equivant seems to have chosen to balance around *predictive parity* instead of *treatment equality*. The following chapter will define the normative ethical theories of utilitarianism, Kantianism and virtue ethics so that they can be later used to analyze the moral impact of forgoing treatment equality.

Ethics

All these statistical equalities are satisfiable; however, as mentioned earlier, some are mutually exclusive. Forcing an algorithm to satisfy predictive parity may compel it to treat protected classes differently. Additionally, this statistical fairness is related to ethical fairness, so before studying the morality and fairness of decision-making algorithms, the normative ethical theories must be defined.

Normative Ethics

Normative ethics is a field of moral philosophy concerned with the criteria and rules of moral right and wrongness (Britannica2016). *Moral theories* are created to answer questions surrounding fundamental moral duties and to define virtuous and vicious character traits. In this research, we will be focusing on three important moral theories: deontology, teleology, and virtue ethics. Deontology

is a moral theory that focuses on the *inherent rightness* of an action in order to establish an ethical standard. Whereas teleological ethics focuses on the *consequence* an action brings. The most common ethical theory in deontology and teleology are *Kantianism* and *consequentialism* respectively. Both of these theories focus on the right and wrong actions of individuals and groups. In contrast to deontology and teleology, *virtue ethics* focuses on individuals and their virtuous and vicious character traits. Instead of an act being either inherently right or right because it increases goodness, virtue ethics consider moral rightness to stem from actions done by virtuous people. Before analyzing algorithmic fairness through the lens of consequentialism, Kantianism and virtue ethics, it is important to understand the core of each moral theory.

Utilitarianism

> "Acts are morally right just because they maximize the amount of goodness in the world" (Moore1903)

Utilitarianism is a prominent consequentialist ethical theory focused on the consequences of actions to justify their moral rightness. In order to find the morally required action, utilitarians believe that the right action maximizes *overall well-being*, and the ends justify the means so long as the ends are good enough. The action that optimizes the ratio of benefits over drawbacks is called the *optimific* choice and is the morally required one. In utilitarianism, wellbeing is the only intrinsically valuable thing and faring-poorly is the only intrinsically bad thing. It states that "an action is morally required just because it does more to improve overall well-being than any other action you could have done in the circumstances" (Shafer-Landau2009).

Kantianism

> "Act only according to that maxim whereby you can at the same time will that it should become a universal law." (Kant1995, p. 33)

Kantianism takes a completely different stance on morality than utilitarianism. It proposes that an action is morally right if, and only if, one's reasons and intentions (*maxim*) behind the action are *universalizable*. A maxim consists of two parts: an action and a reason (e.g. you go to work to help support your family). A maxim is said to be universalizable if, when everyone supports and acts on the maxim, the goal of the action can still be achieved.

The reason for Kant's moral theory to stress universalizability is so that nobody is made an exception. Kantianism believes it to be immoral for one to follow a different rule-set than others, and argues through justness, equality and the ethical concept of fairness instead of optimificity (Shafer-Landau2009). It is, for example, unjust for someone to evade taxes in order to save their sick mother by paying for her treatment, because if everyone would evade taxes, there would be no societal funding for treatment centres, and thus none of the sick could go to treatment.

Using the example above, a utilitarian may argue that it is morally right to evade taxes if the action increases overall well-being. The government may not be able to use the money to bring as much good to the population as this treatment may bring to the mother. Hence, there is a division in morality based on which moral theory one follows. Kantianism finds it crucial to have control over the morality of our actions and even argue that it is impossible to assign credit or blame for things we cannot control. On the other hand, consequentialism argues that credit or blame can only be assigned after the consequences are known, even though these consequences are often not within our control.

Virtue Ethics

> "Moral wisdom is an extremely complicated kind of skill. It does require knowledge of the way the world works, but it demands [...] a great deal of emotional intelligence as well." (Shafer-Landau2009, p. 258)

In contrast with Kantianism and conse-

quentialism, virtue ethics looks at the morality of the agent to find the morality of an action. Virtue ethics is a family of theories that trace all the way back to Aristotle's *Nicomachean Ethics*. While there is no simple set of rules, virtue ethics is based on the principle that in general one should act on virtues and forgo vicious actions.

Virtue ethicists propose that virtues and vices define moral rightness but understand that virtues and vices may conflict in certain situations. When there is a conflict to these simple rules, one requires moral wisdom to understand what action is morally right. They argue that we have to follow the lead of a virtuous person, since it is unlikely for everyone to be morally knowledgeable enough to understand the moral complexities of the actions.

Everyone should aspire to be virtuous, as without it we will not be able to find *eudaimonia*, which can be understood as an intrinsic happiness or flourishing, but in order to become virtuous, one must both be virtuous by nature (character traits) and be morally wise enough to understand the complexities between decisions.

In the example of the sick mother, virtue ethicists may argue that a virtuous person would choose to evade taxes out of virtue (kindness, loyalty and compassion towards their mother) instead of vice (greed) and thus, tax evasion is morally justifiable if it has virtuous intent.

This chapter studied three normative ethical theorems: utilitarianism, Kantianism, and virtue ethics, and highlighted their differences. In order to find the ethical implications of each fairness metric, the next chapter will study the impact of the fairness metrics on the ethical fairness and moral rightness of COMPAS through these ethical theories.

Looking at COMPAS through Normative Ethics

Individual Metrics

Before taking a look at our ethical obligations in balancing treatment equality and predictive parity, let's first study each met-

ric of fairness individually and see why they make sense in the first place. The importance of a decision-making algorithm achieving *overall accuracy equality, statistical parity, conditional procedure accuracy equality* and *predictive parity* can mainly be seen outside of the realm of ethics. These four fairnesses partly dictate how well an algorithm functions as a decision-making tool. First of all, if an algorithm has a lower prediction accuracy for a certain protected group, it is difficult to trust the algorithm when a prediction about said group is made. Thus, a big inequality in overall accuracy is not only ethically questionable, but is also undesirable for the user. Secondly, for statistical parity, if the proportion of correct classifications is different per protected group, it is impossible to not keep the protected group membership in mind during a decision-making process. Thirdly, if conditional procedure accuracy equality is not achieved, then the accuracy of a prediction is dependant on group membership. Even if the overall accuracy of the algorithm is fair, it may be the case that the true negative rate of a protected group is so low that a negative classification of a group member should not be accepted without knowledge of the group membership. Lastly, in the case that predictive parity is not achieved, the accuracy of the algorithms output classes will also depend on protected group membership. That is, if an algorithm has a lower accuracy of classifying one group than another, the output of the more accurately classified group is generally more trustworthy and is on average a better prediction; therefore, not using group membership in a decision will decrease the efficacy of the decisions.

In all these examples, it is paramount that the protected group is recognized within the decision-making process, as failing to include the group membership may result in ill-informed decisions. It is however illegal, by both European and American law, to let a decision-making process depend on protected group membership. This is not only the case digitally, as seen in the General Data Protection Regulation (GDPR), but also in clinical decision-making (eu:gdpr, civilrights1964). To summarize the duality, one is legally prohib-

ited to base any aspect of their decisions on protected group membership, but neglecting group membership may negatively influence the efficacy of an algorithm.

The last form of fairness, *treatment equality*, is heavily linked to predictive parity and prevalence through the impossibility theorem. In contrast with the other four fairness measures, Treatment equality is primarily ethically important. While there are definite ties to Conditional Procedure Accuracy equality, it is important to note that fair treatment of groups is not inherently important for the decision-making process. Assume a situation where two equal sized groups, group A and B, are algorithmically assessed on their meat eating behaviour. In reality, group A, on average, is 70% likely to eat meat, while group B, on average, is 90% likely to eat meat. In this construction, an algorithm that treats group A and B as one group would predict a prevalence of 80%, and misclassify at least 10% of both groups. The only way for the efficacy to increase would be by classifying group A and B partly on their group membership, as this is clearly a statistically significant feature. In reality, even if race is not directly asked by the algorithm, it may find patterns relating to religious beliefs or cultural influences that lay at the root of this prevalence inequality. In large enough data sets, the algorithm may indirectly differentiate on protected groups through a similarity in travel patterns (e.g. to church), living situation, or the prevalence within family and peer group. In this construction protected group membership does not directly affect the predictions and thus does not result in any legal wrongdoing by the users, even though decisions are still indirectly made based on group membership. While an argument can be made in favour of accepting the data as a reality, it is ethically important to understand the intrinsic differences between the groups and how these different prevalence rates came to be.

In the case of COMPAS, we are looking at an algorithm that reasonably achieves Overall Accuracy equality and Predictive Parity. The unfairness of the algorithm stems from the treatment inequality, as black defendants are on regularly classified as significantly riskier than white defendants. This, as we saw in the previous paragraph, stems from the difference in the prevalence of recidivism in the data, as black defendants have a base rate of 0.51, while white defendants have a base rate of 0.39. This difference causes the algorithm to radically classify black defendants as higher risk and as a result misclassifies black defendants as high risk twice as often. The treatment inequality directly influences statistical parity, by increasing the rate of positive classification and decreasing the rate of negative classifications. Additionally, the treatment inequality has an effect on the conditional procedure accuracy equality as it overestimates black defendants risk and thus reduces the true negative rate, while white defendants' risk is underestimated which increases their true negative rate. The inequality of both statistical parity and the conditional procedure accuracy can thus be directly linked to treatment inequality.

In the previous paragraphs, showed that a difference in base rate can lead to unequal treatment even if the group membership is not directly considered. Additionally, the different prevalence rates cause treatment inequality, but in extension also cause an inequality in statistical parity and conditional procedure accuracy equality. In the following subsection, we ethically study the impact of satisfying predictive parity over treatment equality. As we have seen earlier, satisfying predictive parity is important for the algorithm to properly function; however, how should we deal with the treatment inequality that follows from it?

Balancing Predictive Parity and Treatment Equality

Firstly, we need to understand that a fair system does not immediately imply a just system, and vice versa. While the two are often used interchangeably, for our ethical evaluation it is important to study the two definitions. Justice denotes actions that are morally required, while fairness denotes an assessment of whether these actions are morally praiseworthy.

Secondly, it is important that we speak of the tension between public safety and fairness. Maximizing public safety requires de-

taining all defendants, regardless of protected group membership, that are predicted to commit a crime. Satisfying all metrics of fairness would require some group specific thresholds that are often in direct conflict with maximizing public safety. Corbett-Davies (2017) find a real tension between public safety and satisfying the different forms of fairness. They train an algorithm on the Broward County data, optimized to predict cost and increase public safety. Through this algorithm, they estimate an increase of 7% in violent crime if predictive equality, closely related with treatment equality and conditional procedure accuracy equality, was to be satisfied. Thus, while improving the fairness of the system, there is a significant decrease in predicted public safety.

It is from this conclusion that we start looking at the ethical impact of the decision to satisfy predictive parity over treatment equality. We saw earlier that satisfying predictive parity is necessary for a universal algorithm to function. That is, if a risk assessment algorithm is not well calibrated, its classifications cannot be seen independent of group membership, and thus the algorithm cannot be applied equally to all groups. Additionally, we know that attempting to satisfy treatment equality will decrease public safety. This means that in order to treat each group equally, when the prevalence of recidivism differs, we not only directly decrease public safety, we also decrease the efficacy our algorithm and thus misclassify more often, resulting in an even larger decrease of public safety. This leads us to the ethical question: should we significantly decrease public safety in order to treat each group equally and give every individual a fair judgment?

In the following arguments, race will be used in order to speak of protected groups, note however that it does not matter whether we speak of discrimination based on race, gender or religion, as similar arguments can be made for the other groups.

In Virtue Ethics

In order to study the question through virtues and vices, it is important to first expand on the intrinsic differences between each race group and see whether it helps us decide on what choice we are morally obligated to make. In the decision between racial discrimination and public safety, it is necessary to connect racial bias to virtues or vices.

To start, it is important to note that *justice* is a virtue, and that a fundamental principle of justice is 'equals should be treated as equals and unequals should not'. From the difference in the prevalence of recidivism between black and white people one could argue that black and white people should be treated unequally. The flaw in this argument is not obvious, but lies in the simplification of the prevalence of recidivism. Justice dictates that we should treat people with the same standard, which is not the same as treating people the same.

In America, wealth is heavily concentrated among white citizens, and this wealth inequality is rooted in the historic and present unequal treatment of the African-American population. This inequality results in fewer opportunities for black Americans, and is further amplified by employment discrimination, housing segregation, and unequal access to educational opportunities (Asante2016). As a result, America experiences a societal divide between relatively wealthy white communities in which individuals have plenty of opportunities and impoverished communities of color with few vocational or educational opportunities. Lack of opportunities and financial struggle have been proven to positively correlate with crime (Ulmer2012). As a consequence there are more police patrols in underprivileged communities which increases the detection rate of petty crimes (Lum2016). To conclude, the difference in base rate between the two race groups is heavily correlated with opportunity and poverty and is magnified by the increase of police surveillance.

If one is privileged, one generally does not need to break the law to keep themselves and their family alive. Underprivileged people may be put in situations where committing a crime is the only possible way to survive. These crimes may not be completely virtuous; however, they are not strictly vicious either. Thus, crimes committed by privileged people tend to be strictly vicious, while the crimes of under-

privileged people may be virtuous.

Virtue ethics states that we should discriminate based on race group, if and only if, there is a clear difference between the two. We saw that the rate of recidivism of black defendants is higher than that of white defendants. It is incomplete however to not keep in mind the socioeconomic differences between the groups, as black people on average tend to be underprivileged and white people privileged. As highlighted in the previous paragraphs, the difference in crime rate is not explained by race but by difference in wealth and opportunities and crimes of underprivileged people may stem from virtuous intentions. Thus, the belief that we should treat black people differently based on the racial disparity between the prevalence rates is unjust and thus vicious. Instead of increasing public safety by punishing black people for being underprivileged, we should give them more opportunities to reduce their need to commit crimes. A virtuous person would not punish but help the underprivileged, and thus, treatment equality is essential even if it leads to a short term decrease of public safety.

In Kantianism

In the discussion surrounding the use of algorithmic predictions in justice, deontology may seem to be the best match, since both deontology and algorithms are rule-based. While they at first glance seem to match, there is a clear difference in their construction of rules. As we saw in Kantianism, rules must align with a universal form of justice and be universalizable; while in algorithms the rules are artificially created to maximize objectivity, efficacy and statistical fairness. However, the different fairness metrics do not, as shown before, necessarily equate with a universal form of justice. This difference lies at the core of our analysis.

While a developer can personally choose to satisfy treatment equality, Kantianism is brief on the matter. The reasoning of a current day Kantian understanding is independent of COMPAS, as the categorical imperative of Kant coins the equality postulate of universalizablity and equal treatment, and can be extended to a postulate of universal human worth. Thus, a current interpretation of Kantianism can not accept an inequality of treatment regardless of the situation. Note that this analysis is on Kantianism's view on the ethical nature of treatment equality between protected groups, and does not argue about treatment or punishment itself.

In Utilitarianism

Recall that the morally right decision would be the optimific one. If treatment equality is to be justified, we must be able to show that we are better off with it than without it. Treatment equality can be forgone if the societal gain outweighs the individual costs. In the case of race, the question is asked whether it is okay to treat groups differently on the basis of race in order to increase public safety. On the surface, the answer to this seems simplistic, if the consequence of in-equal treatment in the justice system is a safer society then it must increase overall well-being as there is less crime, thus it is our moral obligation to racially discriminate. In practice however, it is not so simple. Other than public safety, we must also consider the societal impact of this systemic racial discrimination.

It is important to understand that the discrimination towards black people in the justice system would lead to a higher incarceration rate of black defendants. As to be expected, the impact of incarceration on families is highly negative, and the family members of incarcerated individuals suffer immensely. The parental incarceration has been shown to potentially lead to emotional and financial strain on the family and to juvenile delinquency (Fishman1981). Additionally, research has shown that a healthy family is negatively correlated with recidivism (LaVigne2005). Thus incarceration of a parent has financial, judicial and emotional effects on families, which should not be disregarded in the consideration of the optimific choice. While this is a brief analysis of the impact of incarceration on social impact, it is immediately observable that the effects are far reaching.

Quantifying the immediate benefits of racial inequality in the justice system seems

relatively easy. In contrast the drawbacks of racial discrimination are more extensive, sometimes subjective, and thus harder to quantify. It is for this reason that there is no immediate clear moral obligation as making up the balance between the benefit and drawback is impossible. This does not immediately matter in the discussion surrounding racial bias and risk assessment algorithms, as utilitarianism does not clearly disregard racial discrimination and thus through consequentialism, an argument can be made to keep unfair risk assessment algorithms. Utilitarianism therefore is able to serve as an ethical disclaimer.

To summarize, by Kant's categorical imperative treatment inequality is unacceptable no matter the situation; through the treatment inequality of COMPAS, the algorithm viciously punishes the underprivileged and thus a virtuous person would have to enforce treatment equality; utilitarianism can be used to argue in favor of treatment inequality through short term reasoning, but also used to argue against it through long term reasoning. To close, it seems our moral obligation to force treatment equality in current day risk assessment even if the base rates of different protected groups differ significantly.

Discussion

The previous analysis focused on our moral obligations concerning treatment equality in judicial algorithms. Even though we cannot immediately extend our findings to all predictive algorithms, as the arguments are partly based on the field the algorithm is used in, we can learn from it and attempt extrapolate to a more general level. It is important to first note that the concept of algorithmic fairness and protected group membership is only applicable in the case of classifying people. One could look at whether an algorithm is biased against wine-gums and favors licorice in a restocking algorithm; however, protected group membership and by extension statistical fairness is inapplicable. Therefore, the next part will only feature decision-making algorithms focused on classification of people.

As seen in the last chapter, Kant's categorical imperative prevents any case of moral inequality to be morally right. Thus, the following analysis will only focus on utilitarianism and virtue ethics. In order to study the consequentialist's stance, we look at the cost and benefit of treatment inequality within each field and try to see if there is a clear optimific choice, whereas through virtue ethics we study the virtues and vices related to the unequal treatment.

Let us start by considering the consequentialist moral obligations of developers surrounding treatment equality. Without going too in depth, let us look at the following algorithms: a college admissibility algorithm; a medical algorithm studying the effectiveness of treatment, and a profiling-based advertising algorithm.

In the case of college admissibility, let us hypothesize a situation where protected group A on average is less admissible than group B due to a combination of a lower grade point average and a less developed CV. Since these two factors are highly influential in the decision-making process, it is easy to predict a lower acceptance rate of group A. Through consequentialism, we study the societal gain in admitting more members of group B and the problems caused by the discrimination of group A. While it is unreasonable to oversee all consequences, as is often the case in consequentialism, the following analysis attempts an approximation of a possible argument.

On the one hand, a short term consequence of discriminating against group A is found in the potential increase of prestige. By increasing their proportion of cum laude graduates, they may attract investors to collaborate on an improved higher education; as higher education should be focused on benefiting society, an improved higher education thus improves overall well-being. On the other hand, one could argue that a higher education that lacks diversity performs worse than one that appreciates differences. Even if we accept a short term benefit for the higher education system, and consequently society, there are long term drawbacks that must be considered. One of the drawbacks can be found in an educational divide between group A and B

which in the long run could lead to a societal underdevelopment of group A. It is impossible to find the optimific choice in such a general scenario; however, as in the case of COMPAS, the optimific choice in this situation seems to be dependent on the handling of short term and long term consequences.

Furthermore, Imagine a medical algorithm that is able to classify two illnesses in patients. Suppose that there is a targeted medication for each class, and different prevalence rates of the illnesses dependent on protected group membership. A wrong prediction, and thus treatment, is not helpful, and even potentially harmful for the patients. Forcing treatment equality in this construction would decrease overall well-being as the algorithm performs worse. This is hurtful for both the protected groups and the society in general. An unequal treatment results in better predictions and thus optimize well-being. Thus, satisfying inequality may be the optimific choice, and thus we are morally obligated to have the algorithm discriminate.

Lastly, suppose an algorithm is used to predict consumer behaviour. Additionally, suppose that the conversion rate of a specific ad is dependent on protected group membership. That is, for example, an ad is 30% effective for group A and 15% effective for group B. Let us assume that seeing this ad or buying the product does not reduce anyone's short or long term overall well-being. Here the advertiser would like to discriminate on group membership because it improves their conversion rate and increases the well-being of the companies employees. Treating group A and B differently would mean advertising more to group A than to group B. In this construction, the optimific choice may be to discriminate on group membership, as it increases the overall well-being of the company and its employees. Additionally, it may improve the well-being of the consumer by offering a much wanted product. In this construction, one clear drawback to targeting based on protected group membership is the loss of potential consumers within group B, which should be balanced by a larger increase of group A consumers. The increase in false negative predictions for group A is notable, however, we assumed that seeing an

ad does not decrease overall well-being. If the company does not discriminate, the benefits are reduced while the number of missed potential consumer is expected to remain consistent (they are just spread out over both groups now). By construction we have found another instance where the optimific choice would be to disregard treatment equality and discriminate on protected group.

In these three constructions, the impact of discrimination changes for the disfavoured group must be noticed. In the case of college admissibility, like in the case of COMPAS, the bias against a protected group hurts the protected group itself and creates a reinforcing loop. Thus, while short term well-being may increase, in the long term, the optimific choice becomes harder to deduce. The medical algorithm favoured both groups and resulted in an overall positive outcome in which case discrimination may actually be desirable. Lastly, in the case of harmless advertisement, neither group benefits nor hurts. Thus, as a rough heuristic, one can look at the impact of the bias against a protected group, if it is beneficial or harmless consequentialism is in favour of forgoing treatment equality, while if it hurts the protected group the benefits to the society need to be substantial, else consequentialism favours treatment equality.

In virtue ethics, it is important to look at the reason behind the inequality. If the inequality in base rate between a protected group A and B stems from, for example, a lack of privilege or opportunity it is vicious to punish the underprivileged and thus treatment equality should be satisfied. However, like in the previously mentioned medical algorithm, if there is a good reason to treat protected groups differently it may be virtuous to do so. In extension it implies that in virtue ethics, there is a possible moral obligation to treat protected groups unequally.

In chapter **??** and the previous part of the discussion, we see that, while a counter argument can be made through utilitarianism, we are morally obligated to fight against the current use of risk assessment algorithms. This fight should probably be fought on European soil as the American political climate has not yet warmed up to data protection regulations;

in 2015 Obama proposed a Consumer Privacy Bill of Rights that got shut down by congress, after which, the topic 'data protection regulation' was removed from the agenda by the Trump administration.

There is reason to believe that the European Union (EU) would be in favour of improving the status quo. In 2018, the GDPR was constructed to restrict the collection and processing of personal data of individuals within the European Union. A possible way for the EU to combat the unethical nature of risk assessment tools is to include them within articles 9 and 22 of the GDPR. The two articles focus on the rights of Europeans by prohibiting the use of protected group membership in algorithmic processes and demands measures to safeguard the subject's rights in algorithmic decision making (Articles 9 & 22 of eu:gdpr). A combination of Article 22(4) and 9(2)(g) states: "Decisions shall not be based on special categories of personal data [...], unless processing is necessary for reasons of substantial public interest [...]." (Articles 9(2)(g) & 22(4) of eu:gdpr). Decisions surrounding criminal convictions and offences are currently excluded from this regulation, including them in article 9 and 22 forces the developers of risk assessment algorithms to satisfy treatment equality in order for their algorithm to be in the public's interest. Note that this line of reasoning only holds if one accepts that acting on our moral obligation is in the public's interest.

To close, even when treatment equality is satisfied and the algorithm has an acceptable efficacy, we may not want to use predictive tools in every situation. We saw in the introductory chapter, when algorithms are used as a tool in decision-making the transparency of the decisions decrease significantly. Thus, a fair and well-performing risk assessment tool may make a court sentence unexplainable and goes against our current western belief of transparent justice.

2   Conclusion

The primary contribution of this thesis was to ethically study the complications surround-

ing algorithmic decision-making, specifically fairness. The preliminary analysis focused on COMPAS data placed in Berk's model of statistical fairness, and discussed each fairness metric through its effect on COMPAS. The primary analysis studied the practical necessity of each fairness metric and discussed treatment equality, and our moral obligation to satisfy it, through the lens of three normative ethical approaches, namely consequentialism, Kantianism and virtue ethics. In the case of COMPAS, Kantianism and virtue ethics heavily oppose unequal treatment. The former argues, through its categorical imperative, against any unequal treatment, while the latter argued the vicious nature of punishing an underprivileged individual for their lack of privilege. Consequentialism can be interpreted to be less decisive about the moral obligation of satisfying treatment inequality. One may argue that an optimific choice against treatment equality is made when the short term benefits are larger than the short term drawbacks, which is the case for COMPAS; however, one can also argue that in the long run the continuation of treatment inequality only increases the racial divide causing many to suffer. Noteworthy is the fact that, while Kantianism and virtue ethics heavily oppose treatment in equality, there exist constructions that allow consequentialist reasoning to morally obligate an unequal treatment in the justice system.

The latter conclusions may change when one scrutinizes decision-making tools that are used outside the judicial realm, such as, algorithms used in the private sector. Kant's categorical imperative supposes that every individual should be treated by the same standard. Virtue ethics' view can be summarized in a few cases: first of all, if all protected groups benefit from unequal treatment, there is a moral obligation to do so; secondly, if an inequality in base rate is found through lack of privilege or opportunity, it is vicious to punish the underprivileged group for it, thus treatment equality should be satisfied; lastly, all other constructions require further study. Consequentialism requires one to look at the societal impact of enforcing a protected group bias to decide one's moral obligation. If the dis-

crimination is beneficial or harmless for a protected group consequentialism one can argue that forgoing treatment equality is morally required, while if it hurts the protected group the benefits to the society need to be substantial, else consequentialism favors treatment equality.

To conclude, we are morally obligated to further study and improve current risk assessment tools and should strive for an equal treatment of protected classes. It is important to carry out further research in the fields of algorithmic opacity, efficacy, and fairness, as predictive algorithms are already widely used and the discussion surrounding their limitations seem to be just warming up. There is a tendency among software engineers to speak of the fairness and rightness of their algorithm, while in the same breath mention their ethical ignorance. The current analysis has shown that the statistical fairness metrics, do not always match the ethical ideas about fairness and justice. By ignoring our moral obligation and accepting ignorance, developers may perpetuate and create deeply rooted inequalities in society that will be intertwined with the future's data dependency.

Works Cited

Angwin, J., Larson, J., Mattu, S., Kirchner, L. (2016a). How we ana- lyzed the compas recidivism algorithm. *ProPublica*. Retrieved 2017-04-30, from

Angwin, J., Larson, J., Mattu, S., Kirchner, L. (2016b). Machine bias: there's software used across the country to predict future criminals. and it's biased against blacks. *ProPublica*. Retrieved 2017-04-30, from

Asante-Muhammed, D., Collins, C., Hoxie, J., Nieves, E. (2016). The Ever-Growing Gap. Retrieved from

Athey, S., Katz, L. F., Krueger, A. B., Levitt, S., Poterba, J. (2007, May). What does per- formance in graduate school predict? graduate economics education and student outcomes. *American Economic Review*, 97(2), 512-520. doi: 10.1257/aer.97.2.512

Barocas, S., Selbst, A. (2016). Big Data's Disparate Impact. *California law review*, 104(1), 671–729. doi: 10.15779/Z38BG31

Berk, R. (2012). *Criminal justice forecasts of risk a machine learning approach*. New York, NY: Springer.

Berk, R., Heidari, H., Jabbari, S., Kearns, M., Roth, A. (2017). Fairness in Criminal Justice Risk Assessments: The State of the Art. , 1–42. Retrieved from

Bunge, M. (1963). A general black box theory. *Philosophy of Science*, 30(4), 346– 358. Retrieved from

Chalfin, A., Danieli, O., Hillis, A., Jelveh, Z., Luca, M., Ludwig, J., Mullainathan., S. (2016). Productivity and Selection of Hu- man Capital with Machine Learning. *American Economic Review*, 106, 124–127.

Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data*, 5(2), 153–163. doi: 10.1089/big.2016.0047

Civil Rights Act of the United States of America. (1964). *Civil rights act*. Pub.L. 88-352,78 Stat. 241.

Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., Huq, A. (2017). Algorithmic decision making and the cost of fairness. *CoRR*. Retrieved from

Corbett-Davies, S., Pierson, E., Feller, A., Goel, S. (2016). A computer pro- gram used for bail and sentencing deci- sions was labeled biased against blacks. it's actually not that clear. *Washington Post*. Retrieved 2016-10-17, from

Council of the European Union. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*. Retrieved from

Dawes, R. M., Faust, D., Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, 243(4899), 1668–1674. Retrieved from

Dieterich, W., Mendoza, C., Brennan, T. (2016). COMPAS Risk Scales: Demon- strating Accuracy Equity and Predic- tive Parity. Retrieved from

Dressel, J., Farid, H. (2018). The accu- racy, fairness, and limits of predicting recidivism. *Science Advances*, 4(1), 1–6. doi: 10.1126/sciadv.aao5580

Dwork, C., Hardt, M., Pitassi, T., Reingold, O., Zemel, R. (2011). Fairness Through Awareness. Retrieved from doi: 10.1145/2090236 .2090255

Encyclopaedia Britannica. (2016, Apr). *Normative ethics*. Retrieved from

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542, 115–. doi:http://dx.doi.org/10 .1038/nature21056

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York, NY, USA: St. Martin's Press, Inc.

Fishman, S. H. (1981). Losing a loved one to incarceration: The effects of imprisonment on family members. *The Personnel and Guidance Journal*, 371–375. doi: http://dx.doi .org/10.1002/j.2164-4918.1981.tb00572.x

Flores, A., Lowenkamp, C. T., Bechtel, K. (2016). False Positives, False Negatives, and False Analyses: A Rejoinder to "Machine Bias". (August), 1–33.

Geraghty, K. A., Woodhams, J. (2015). The predictive validity of risk assessment tools for female o enders: A systematic review. *Aggression and Violent Behavior*, 21, 25– 38. Retrieved from http://dx.doi.org/10 .1016/j.avb.2015.01.002 doi: 10.1016/ j.avb.2015.01.002

Guy, L. S., Packer, I. K., Warnken, W. (2012). Assessing Risk of Violence Using Structured Professional Judgment Guidelines. *Journal of Forensic Psychology Practice*, 12(3), 270– 283. doi: 10.1080/15228932.2012.674471

Hardt, M., Price, E., Srebro, N. (2016). Equal- ity of Opportunity in Supervised Learning. , 1–22. Retrieved from

Jaconelli, J. (2002). *Open justice a critique of the public trial*. Oxford: Oxford University Press.

James, G., Witten, D., Hastie, T., Tibshirani, R. (2014). *An introduction to statistical learning: With applications in r*. Springer Publishing Company, Incorporated.

Kant, I. (1995). *Kant: Ethical philosophy: Grounding for the metaphysics of morals, and, metaphysical principles of virtue, with, "on a supposed right to lie because of phil- anthropic concerns"* (J. W. Ellington, Trans.). Hackett Publishing Company.

Kleinberg, J., Lakkaraju, H., Leskovec, J., Lud- wig, J., Mullainathan, S. (2017). Human Decisions and Machine Predictions. *The Quarterly Journal of Economics* (January), 237–293. doi: 10.1093/qje/qjx032.Advance

Kleinberg, J., Mullainathan, S., Raghavan, M. (2016). Inherent Trade-O s in the Fair Determination of Risk Scores. , 1–23. Retrieved from *http://arxiv.org/abs/1609 .05807 doi: 10.1111/j.1740-9713.2017 .01012.x*

La Vigne, N. G., Naser, R. L., Brooks, L. E., Castro, J. L. (2005). Examining the e ect of incarceration and in-prison family contact on prisoners' family relationships. *Journal of Contemporary Criminal Justice*, 21(4), 314– 335. doi: 10.1177/1043986205281727

Latour, B. (1999). *Pandora's hope: essays on the reality of science studies*. Cambridge, MA: Harvard University Press.

Lum, K., Isaac, W. (2016). To predict and serve? *Significance*, 13(5), 14–19. doi: 10.1111/j.1740-9713.2016.00960.x

Meehl, P. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. University of Minnesota Press.

Moore, G. E. (1903). Principia ethica. Camebridge University Press.

Northpointe. (2015). Parctitioner's Guide to COMPAS Core.

Northpointe Inc. (2011). *CORE Questionnaire Risk Assessment.* Retrieved from

Ohlin, L., Dudley, D. (1949). The Efficiency of Prediction in Criminology. *American Journal of Sociology*, vol. 54. pg 441– 452.

O'Neil, C. (2016a). How algorithms rule our working lives. *The Guardian*. Retrieved 2017-04-30, from

O'Neil, C. (2016b). *Weapons of math destruc- tion: How big data increases inequality and threatens democracy*. New York, NY, USA: Crown Publishing Group.

Ramo, R. C., Black, A. D., President, V., Laycock, D., President, V., Marshall, M. H., Friedman, P. L. (2014). Model Penal Code : Sentencing APPENDIX B Black Letter of Ten- tative Draft No . 3 The American Law Insti- tute. (3).

Reinsel, D., Gantz, J., Rydning, J. (2017).

Data Age 2025: The Evolution of Data to Life-Critical. *IDC White Paper; Sponsored by Seagate* (April), 1–25. Retrieved from

Shafer-Landau, R. (2009). *The fundamentals of ethics*. Oxford University Press.

Skeem, J. L., Lowenkamp, C. T. (2016). Risk, Race, and Recidivism: Predictive Bias and Disparate Impact*. *Criminology*, 54(4), 680– 712. doi: 10.1111/1745-9125.12123

Starr, S. (2014). Evidence-Based Sentencing and the Scientific Rationalization of Discrim- ination. *Stanford Law Review*, 66.

Ulmer, J. T., Harris, C. T., Ste ensmeier, D. (2012, September). Racial and ethnic disparities in structural disadvantage and crime: White, black, and hispanic comparisons*. *Social Science Quarterly*, 93(3), 799–819.

Yang, M., Wong, S. C. P., Coid, J. (2010, September). The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *American Psychological Association*, 136(5), 740–767.

Sciences

The Impact of a Transition to Autonomous Vehicles on Total Energy Consumption in the Passenger Vehicle Transport Sector in the Netherlands

Julian Visser

Abstract

With new autonomous driving technologies being developed by many car manufacturers, an often unanswered question remains what the impact on the total energy consumption of such a transition would be. Using the Netherlands as a case study, this study researches the effects on the total energy consumption of passenger vehicles in the Netherlands, if a transition to autonomous electric vehicles (AEV) were realized. Note that this study focuses on a double transition, from fossil fuel to electric powered and man controlled vehicles to autonomous vehicles, which both affect the energy consumption differently. Results were gathered through modeling the energy efficiency effects using the available data from the Dutch passenger vehicle transport sector. It was found that total energy consumption will decrease 5.1% or even up to 57.2% in a best-case scenario. However, if only considering the effects of a transition to autonomous vehicles (thus leaving out the electrification of vehicles), energy consumption can either increase 23.2% (worst-case scenario) or decrease 28.2% (best-case scenario), showing the uncertainty of the impact of a transition to autonomous driving.

Introduction

In recent years, the use of Artificial Intelligence (AI) has developed rapidly, including in the vehicle industry. Vehicle manufacturers, such as Tesla and BMW, are already implementing AI into our everyday vehicles in the form of lane assist and automated parking. Behind the scenes, companies are working hard towards putting more AI technology in our vehicles to simplify and, eventually, achieve completely autonomous driving. With new intelligent assists coming out every year, we will most likely ultimately reach autonomous vehicles (AV). Over the past decade, energy consumption of the passenger vehicle transport sector in the Netherlands has been slowly declining (CBS 2017d); however, on a global scale energy consumption of the transport sector is on the rise (Conti et al. 2016)(Herzog et al. 2005)(Bashmakov et al. 2014). With climate goals set, energy efficiency increases and energy savings have become important agenda items for political programs.

In the literature, the focus of studies on AV is often focused on their driving behaviour compared to that of vehicles controlled by humans. Instead, this paper addresses the differences in energy consumption of Autonomous Electric Vehicles (AEV) and vehicles controlled by humans.

To narrow down the scope of the research, this study compares human controlled and AEV, as if they co-existed on the same road, making use of the same infrastructure. It must be said that AV perform better with increasing numbers, as it is possible for AV to communicate with one another,

thus increasing traffic fluidity considerably (Meyer et al. 2017). However, studies on these systems are limited and remain mainly theoretical, as these AV have not yet been tested on a larger scale. Furthermore, this study only considers passenger vehicles, thus excluding heavier transport, such as trucks. The study thus answers the following question: What will be the impact of a transition to electric autonomous passenger vehicles on the energy consumption in the transport sector of the Netherlands?

To answer this question a model has been developed to calculate the impact of AEV on energy consumption for different scenarios. To find the difference in energy consumption between AEV and vehicles controlled by humans, where the main focus is on driving behaviour differences between AI and humans, which is determined by the areas of vehicle movement (entailing air resistance, accelerating/braking, AI power consumption and rolling resistance) and navigating. Further implications of AEV on energy consumption, such as their influence on vehicle accidents, vehicle aesthetic developments and eventually infrastructure development differences, will also be discussed shortly.

Research context

An elaborate overview of current studies on energy efficiency increases and decreases due to autonomous driving is given by Wadud et al. (2016). In this overview studies researching energy efficiency increases due to a transition to AV were

compiled. The various fields discussed are: congestion mitigation, automated eco-driving, platooning, improved crash avoidance and what they call "right-sizing" of vehicles. Regarding energy efficiency decreases due to a transition to AV, the paper mentions one possible increase: the total travel time by passenger vehicle, which is caused by both a reduction in time spent driving (i.e. making driving less of hassle) and possible new user groups of these AV (ie people that cannot currently drive but would be able to with the introduction of AV). Lastly, the paper covers potential fuel mix changes, suggesting that a transition to AV could overcome some of the hurdles of alternative fuels, thus leading to a reduction in fuel emissions intensity. However, they leave the quantitative analysis of this topic to further research.

This study differs from Wadud et al. (2016), in its focus on the Netherlands by providing a figure on both individual and national scale in either energy consumption increase or decrease of a transition to AEV. The various subsections mentioned in the introduction together with the areas of energy consumption mentioned by Wadud et al., will provide guidance material to the research. Further research on these subsections is discussed below.

An interesting model for the energy consumption of a moving vehicle is given by Mackay (2009). According to Mackay there are three energy consumers for vehicle movement. To simplify his model of the energy consumption for vehicle movement, he claims that there are two different regimes during driving, the "Accelerating and Braking" regime and the "Air Resistance" regime. During short driving distances the "Accelerating and Braking" is the dominant consumer of energy, while for longer stretches with less regular braking aerodynamic drag becomes increasingly important. Mackay calculates that for some (by him defined, see Appendix A.1) average vehicle the Accelerating and Braking regime becomes the dominant energy consumer when the braking distance is less than 750 meters, while the Air Resistance regime is the dominant regime for braking distances above this value. Besides the two regimes, Mackay identifies a third energy consumer of vehicle movement as the rolling resistance, which is always present and depends on your speed. The combination of the formulas describing energy consumption of the two regimes and rolling resistance, provide a relatively accurate model of the total energy consump-

tion of a moving vehicle.

The formulas discussed by Mackay (2009) are used as a basis for the model in this study. These formulas will be further discussed in the methodology. However, first more research regarding the energy consumption of both AV and conventional vehicles will be discussed, as well as research covering decreases in energy consumption due to AEV.

He et al. (2012) calculated energy efficiency increases in AV in urban areas as a result of an energy management system and a cycle optimization algorithm. Using predictive traffic data, this algorithm managed to increase fuel efficiency by 56-86% for conventional vehicles. Applying this algorithm to a plug-in hybrid electric vehicle boosts fuel efficiency to roughly 115% due to the plug-in hybrid electric power management system. Translating fuel efficiency to energy consumption, we find that autonomous conventional vehicles use 35.8% to 46% less energy than their human controlled counterparts. For plug-in hybrid electric vehicles we find that this value is 53.5%. Comparing these values, it can be concluded that an AEV (in this study semi-electric) has additional benefits besides autonomous driving. Therefore, we must note that the transition to electric vehicles also has an impact on energy consumption within the transport sector. Further research on this topic has been done by Ambel (2017), which gives an extensive overview of energy efficiencies of electrical engines, internal combustion engines and fuel cells, also entailing the refinement of the fuel source. In this study it is shown that electric vehicles, powered exclusively by renewable energy, are far more efficient than the other two possibilities (similar efficiencies are concluded in Yazdanie et al. (2014)). However, the current Dutch electricity production is not exclusively based on renewable energy sources. Research on the electricity composition in the Netherlands is done yearly by Energie Onderzoek Nederland (Energy Research the Netherlands, ECN) and Centraal Bureau voor de Statistiek (Statistics Netherlands, CBS), ECN's most recent report is Schoots et. al (2017), while CBS can provide data from 2015 (CBS 2017c). Furthermore, Ambel (2017) assumes that electricity consumed for the refinement of the other fuel sources is also exclusively generated by renewable energy sources. These assumptions are important to note and are taken into account whenever figures from this research are used in the study.

Furthermore, Mensing et al. (2013) showed, through research similar to He et al. (2012), that a driving algorithm, which they call eco-driving, is able to decrease energy consumption up to 34% in urban areas. However, they argued that eco-driving does not take safety for the passenger into consideration. Therefore, the actual decrease in energy consumption in everyday use lies in the range of 15% to 28%.

As was mentioned by Wadud et al. (2016), slip-streaming (platooning) could have significant impact on fuel consumption and thus energy consumption. Duan et al. (2007) studied the effect of slipstreaming on Californian highways by using both an experimental setup and a model. During the research they varied the distance between the slipstreaming vehicles and the type of vehicles. Experimental results showed a 61% drag reduction for the second vehicle, while the model showed 40-60% drag reduction (again for second vehicle) due to platooning. Similarly, Wadud et al. (2016) discusses a decrease from 45% up to 55% for highway driving. Comparable research has been done by Zhu and Yang (2011). This study used computational fluid dynamics to simulate the flow field over two generic sedans and then compared this to the flow field of a single sedan. Based on the difference in flow fields, comparisons on the aerodynamic drag were done, resulting in a clear relation between the distance separating the two vehicles and the drag coefficient of the second vehicle. Furthermore, platooning also showed effects on the drag coefficient of the leading sedan, as its rear-end air flow was disturbed by the closely following second sedan. Nevertheless, this effect was significantly smaller than the decrease in drag coefficient for the second sedan, leading in an overall drag coefficient reduction due to platooning. This effect was around 10% if only two sedans were participating in the platoon. However, in an urban setting slightly different values were found by Zabat et al. (1995). They found that, due to platooning, on average vehicle energy consumption could be lowered 5-10% in an urban setting.

Research regarding rolling resistance of passenger vehicles has already been done by a variety of institutes and studies. In 'Fundamentals of Vehicle Dynamics' Gillespie (1992) devotes a short chapter to rolling resistance, where he mentions that on average the rolling resistance coefficient is around 0.015 for passenger vehicles on con-

crete roads. Furthermore, Gillespie discusses energy consumption in the two regimes very similar to Mackay (2009), thus further supporting the model set up by Mackay. More data on rolling resistance for passenger vehicles has been collected by Société de Technologie Michelin (2003). Results from this study were comparable to values found in Gillespie (1992). Additionally, the Société de Technologie Michelin (2003) concludes that the rolling resistance stays constant up to 120 km/h.

Another possible decrease in energy consumption of AV was presented in the form eco-routing by Boriboonsomsin et al. (2012). Eco-routing is an optimization algorithm and suggests to the driver the route that uses the least fuel. The study found that eco-routing could translate into a 13% decrease in fuel consumption. However, there is a major drawback to eco-routing: choosing the optimal route fuel wise often leads to concessions on the time spent driving. Therefore, the decrease in fuel consumption comes accompanied with an increase in travel time of 5-20% (Meyer et al. 2017).

Taking a step back and looking at research focusing on direct effects of a transition to AV, we find that Meyer et al. (2017) provides a thorough research on traffic fluidity impacts of such a transition. According to this study, AV will increase accessibilities significantly as traffic fluidity increases. However, an increase in accessibility will also increase travel demand (Hills 1996), counteracting decreases in total energy consumption. Nevertheless, the positive impact of traffic fluidity overrules the negative effects of higher travel demands. An additional positive effect of a transition to AV on traffic fluidity not mentioned by Meyer et al., is the decrease in vehicle accidents. According to NHTSA (2008) up to 90% of the accidents in the United States are caused by human errors. Research done by Bertoncello and Wee (2015) showed that after a complete transition to AV up to 190 billion dollars could be saved for the United States economy alone.

Scope and limitations

The scope of the study is to quantify the impact of a transition to AEV on the energy consumption in the passenger vehicle transport sector in the Netherlands. As has become clear in chapter 2, such a transition has two elements that in-

fluence energy consumption: the transition to autonomous driving and the transition to electric vehicles from conventional vehicles. Therefore, the research question should be split into two parts: the transition to autonomous vehicles and the transition to electric. After separating, the effects of these transitions are calculated for a single vehicle. Finally, the separated effects are combined, and the impact on a national scale is calculated. Both of these questions require a variety of other questions to be answered first.

- What is the average difference in energy consumption between a single autonomous passenger vehicle and a single conventional passenger vehicle in the Netherlands?

  - What is the effect of platooning on the drag coefficient of AV?

  - How does autonomous driving influence energy consumption during accelerating and braking?

  - What is the average speed of vehicles in urban and non-urban areas?

  - What is the average braking distance of vehicles in urban and non-urban areas?

  - What is the average energy consumption of an average vehicle?

  - What is the power consumption of AI driving a vehicle?

- What is the average difference in energy consumption between a single electric passenger vehicle and a single conventional passenger vehicle in the Netherlands?

  - What is the average motor efficiency of EV and conventional vehicles in the Netherlands?

  - What is fuel production efficiency for petrol?

  - What is the electricity mix in the Netherlands?

A number of limitations exist concerning this study. Although CBS can provide a wide variety of data, speed averages in urban areas and non-urban areas are unknown. Similarly, braking distances for these two areas are also unknown. Therefore,

this study takes averages based on the earlier explained ground rules. If these values differ significantly from real values, results from this study become unsuitable for the Netherlands.

Furthermore, this study will not take any increases or decreases in vehicle use due to AV (Meyer et. al 2017) into consideration. Additionally, it will also not acknowledge any other benefits from autonomous driving, nor will it recognize any pros or cons of human-driving. For example, it is theorized that due to wireless interaction among AV, traffic fluidity can become even further increased than what has been suggested in this paper. Traffic lights, roundabouts, intersections and speed bumps could become obsolete. However, some participants (cyclists, pedestrians, etc.) are not able to communicate with these AV, so this raises new problems regarding infrastructure. Most predictions therefore remain speculative and hard to quantify. Thus, in the trend of not acknowledging any other benefits of a transition to AV, energy consumption influences from infrastructure changes will not be considered. In other words, this study focuses solely on the energy consumption of conventional vehicles and AEV in the Netherlands. Any societal changes (and their accompanied energy consumption in/decreases) from a transition to AEV are left for other studies to analyze. Nevertheless, it should be noted that even if a transition to AEV causes major societal changes in passenger vehicle use, results from this study regarding individual energy consumption of AEV are still be applicable.

Relevance

The transportation sector makes up 29% of total world energy consumption (EIA 2018) (see pie chart below), of which roughly 50% can be ascribed to passenger vehicles (Conti et al. 2016). So, global energy consumption consists of a significant part of passenger vehicles, meaning that any energy efficiency increases could have major effects. Additionally, environmental scientists have suggested energy efficiency increases in a variety of sectors, preferably major energy consuming sectors, to meet the goals stated in the 2015 Paris climate agreement (Bashmakov et al. 2014)(Pachauri et al. 2014). Considering the hefty energy consumption of passenger vehicles in the transportation sector, the sector applies well for potential

energy efficiency increases. Therefore, this study looks into the effect on energy consumption of a transition to AEV. From this study one could either find reason to stimulate AEV or continue research into different areas of energy efficiency increases in the transportation sector. Therefore, this study provides the figures to help make the right decisions towards meeting the 2015 Paris climate agreement in the Netherlands. Furthermore, this study could lay the foundation for various other studies in other countries concerning their transition to AEV, as this study will thoroughly describe the aforementioned model and will consider a wide range of possible scenarios dependent on the level of urbanization, average vehicle speeds and average braking distance. Although results from this study may not be directly applicable to other countries, the results will nevertheless provide a good indicator of the effectiveness on energy consumption of a transition to AEV.

Methodology

Literature review

As was clarified in chapter 2, the energy consumption of a passenger vehicle depends on a wide variety of components. Data concerning the energy consumption of both AEV and human controlled vehicles will be needed to eventually make an accurate calculation. Details on traffic in the Netherlands is mostly be provided by CBS, which is deemed to be a highly reliable source. Any other claims concerning differences in energy efficiency between AEV and human controlled vehicles are supported by the available literature. Most literature was acquired through CataloguePlus (the database of the University of Amsterdam). A summation of some the keywords for finding current research in this field were: *Autonomous, Self-driving, Efficiency, Cars, Vehicles, Platooning, Eco-driving, Eco-routing, Electric, Plug-in Hybrid, Transition, Artificial Intelligence and Rolling resistance*. Lastly, snowballing techniques were used to efficiently get more in-depth papers for specific details and values.

The Model

To answer the research question a model is created, based on a model for vehicle movement set up by Mackay (2009) and adjusted for AEV whenever needed. The figure below shows the basic steps that are required to calculate the energy consumption of AEV in the various regimes for a specific braking distance. Below the figure the formulas needed for the calculations are explained, starting with the separate parts that have to do with the energy consumption of a car, which are used to calculate the average speeds in urban and non-urban areas.
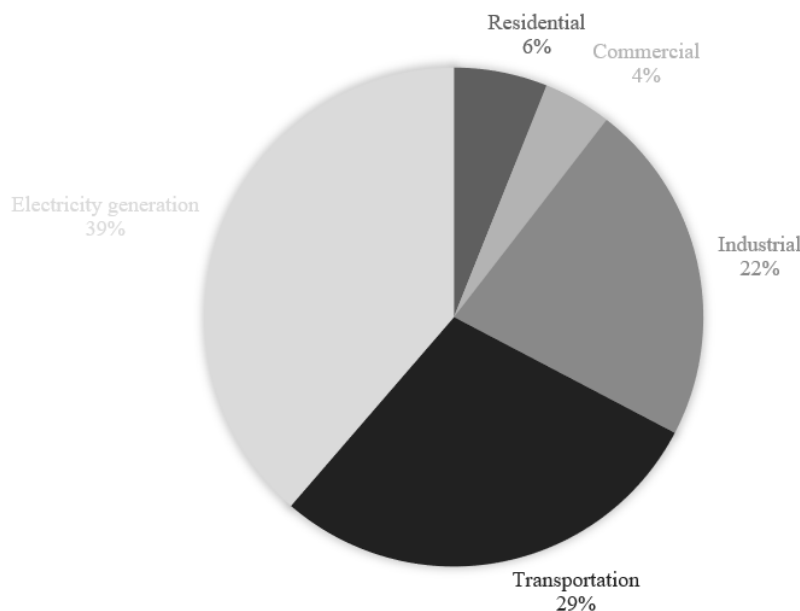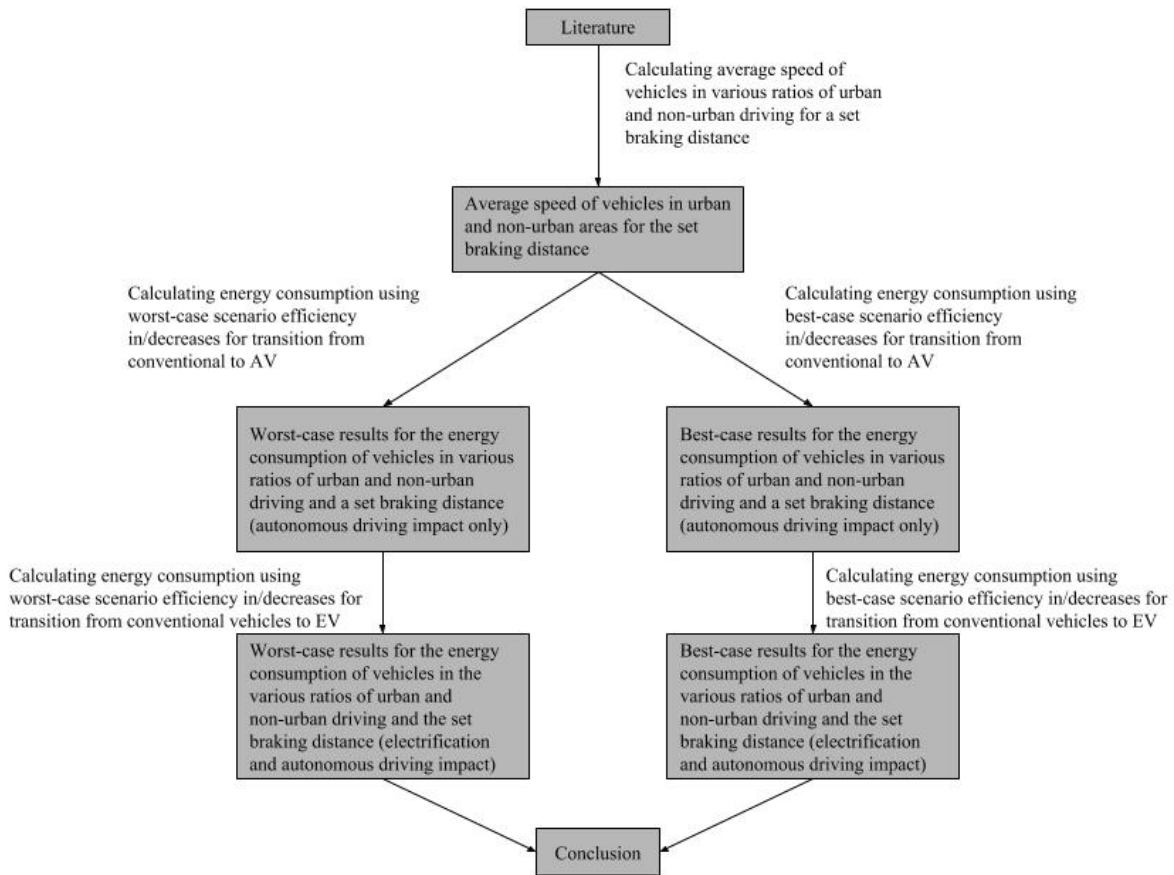

Figure 1: Primary energy consumption by sector (Based on EIA (2018))

Steps to calculating energy consumption for AEV for a specific braking distance

Accelerating and Braking regime

Imagine a driver driving in a city, where he constantly needs to accelerate and brake due to traffic lights. Every time the driver is halted by a traffic light he hits the brakes and converses all his kinetic energy into heat in his brakes. Depending on the distance between the traffic lights and his or her speed, he or she has to hit his brakes after every x amount of seconds. Therefore, the driver consumes energy at the following rate:

$$\text{Power}_{accelerating\ and\ braking} = \frac{E_k}{\text{time between braking events}}$$
(1)

where E_k is the kinetic energy of the vehicle, m_c is the mass of the vehicle, d the distance between braking events and v the speed of the vehicle. However, this formula does not take regenerative braking into consideration, which is a stan-

dard in EV. The efficiency of regenerative braking is about fifty percent, and thus halves the energy lost during braking:

$$\text{Power}_{accelerating\ and\ braking\ AEV} = \frac{m_c v^3}{4d}.$$
(2)

Air Resistance regime

While driving through the city, the vehicle also consumes energy by plowing through the air. As the vehicle moves, it is essentially moving air from one place to another, thus giving the air kinetic energy. The kinetic energy of moving air can be described by:

$$\text{Kinetic energy of moving air} = \frac{m_{air} v^2}{2}$$
(3)

m_{air} can be rewritten by imagining that the vehicle creates a vehicle shaped tube in the air it has driven through. This tube of air can be described

by $\rho c_d A_{vehicle} v t$ (the density of the air, drag coefficient of the vehicle, area of the vehicle, speed of the vehicle and time driven respectively). To then energy consumption of air resistance we simply divide by the time, finding

$$\text{Power}_{air\ resistance} = \frac{m_{air}v^2}{2} = \frac{\rho c_d A_{vehicle} v t v^2}{2} = \frac{\rho c_d A_{vehicle} v^3}{2} \quad (4)$$

Rolling Resistance

Mackay discusses rolling resistance as the final energy consumer of vehicle movement. Calculations with the rolling resistance combine all rolling inefficiencies of a car in to C_{rr}. With this value the energy consumption is calculated to be:

$$\text{Power}_{rolling\ resitance} = C_{rr}m_c gv. \quad (5)$$

Total Power Consumption of Vehicle Movement

Combining formulas (2), (4) and (5), the total energy consumed to move a human controlled vehicle is given by

$$\text{Power}_{total} = \frac{\rho c_d A_{vehicle} v^3}{2} + \frac{m_c v^3}{2d} + C_{rr}m_c gv. \quad (6)$$

Note that this formula needs to be adjusted for the efficiency of the motor. Furthermore, this formula also needs to be modified to be applicable to AEV, as the A.I. driving the vehicle must be taken into consideration. So for an AEV we find:

$$\text{Power}_{total\ AEV} = \frac{\rho c_d A_{vehicle} v^3}{2} + \frac{m_c v^3}{4d} + C_{rr}m_c gv + \text{A.I. power consumption.} \quad (7)$$

The increase in energy consumption due to the AI power consumption of driving an AEV is significant. Current autonomous prototype vehicles approximately consume 2000 Watts, compared to human drivers, who use roughly 20 Watts, this number is enormous (Welling 2017)[1].

Setting up the Model Scenarios

Many areas where AEV have a higher energy efficiency than conventional vehicles were discussed in chapter 2. Most of these efficiency increases apply to one of the driving regimes or a specific area of driving, where the difference seems to lie between urban and non-urban driving. Average driving speed differs enormously in these two areas, which influences energy consumption in the specific driving areas. Therefore, the model separates energy consumption in urban and non-urban areas by accounting for two different average speeds. Adjusting formula (6) accordingly, it transforms into:

$$\text{Power} = \frac{\rho c_d A_{vehicle} v_{non-urban}^3}{2} + \frac{\rho c_d A_{vehicle} v_{urban}^3}{2} + \frac{m_c v_{urban}^3}{2d}$$
$$\frac{m_c v_{non-urban}^3}{2d} + C_{rr}m_c g(v_{urban} + v_{non-urban}). \quad (8)$$

Similarly adjusting formula (7) leads to

$$\text{Power AEV} = \frac{\rho c_d A_{vehicle} v_{non-urban}^3}{2} + \frac{\rho c_d A_{vehicle} v_{urban}^3}{2} + \frac{m_c}{2}$$
$$\frac{m_c v_{non-urban}^3}{4d} + C_{rr}m_c g(v_{non-urban} + v_{urban}) + \text{A.I. power consu}$$

To calculate energy consumption of *one* passenger vehicle using (8), the braking distance and average speed (in both urban and non-urban areas) are needed. However, these variables are unknown for the Netherlands. Luckily, CBS has done research on travel time, travel distance, vehicle weight, total energy consumed by all passenger vehicles, and the total amount of passenger vehicles (CPS 2016c)(CBS 2017d)(CBS2017a)(CBS2016a), making it possible to calculate average power consumption for a conventional vehicle (Appendix A.2). For consumption solution for the *speed* in urban and non-urban areas can be calculated through (9) using two variables: the *distance* covered in urban and non-urban areas and the *braking distance* in urban and non-urban areas (all other variables are retrievable through sources or are calculated in Appendix A). Both these variables are unknown for the Netherlands, therefore the model calculates average speeds for 5 different ratios of distance covered in and outside urban areas. The ratios are 1:1, 5:1, 1:5, 3:2, 2:3, based on an average travel distance of about 32 kilometers (Appendix A.4). So for the ratio 5:1, 26.7 kilometers were driven in ur-

---

[1]Although this figure comes from a respectable Dutch newspaper, the actual source is from a less respected magazine named Wired. This source does not provide the origin of the figure

ban areas, while 5.3 kilometers were traveled in non-urban areas. Furthermore, each ratio is analyzed for a set of braking distances, consisting of 100, 200, 300, 359, 400, 500 and 700 meters. The braking distances are chosen based on rounded average distances found between crossings, roundabouts and traffic lights in Amsterdam. The values stop at 700 meters, as for higher values for braking distance, energy losses due to braking become negligible. The urban vs non-urban ratios are determined based on a variety of travel routes in the Netherlands. Combining the ratios and braking distances with the average power consumption per vehicle (See appendix B for the various calculations), formula (8) provides a range of combinations of time spent in and outside urban areas, and thus average speed in these areas. To determine the right combination of time spent in urban and non-urban areas, the total travel time must be considered. The two times found with formula (8) must add up to a total travel time of 2484.6 seconds (see appendix A.5), limiting the range of results.

Applying Efficiency Changes to the Model

After artificially generating average speeds for urban and non-urban areas in the Netherlands, formula (9) is used to calculate the energy consumption of AEV. This is where the energy efficiencies discussed in the research context are applied. Depending on the area of impact of the efficiency increase, the model is altered. Additionally the mass of the car is altered, as electric vehicles are slightly heavier than their Internal Combustion (IC) counterparts (See Appendix B). First, a worst-case scenario is considered, where only the lowest energy efficiency increases are applied and no eco-routing is considered. Thereafter, a best-case scenario is calculated where the model is altered using the best energy efficiency increases and eco-routing. A list of energy efficiency changes for both scenarios is given below (values are based on researches discussed in the research context). Note that for the calculations of the efficiency of the electric and IC engine energy losses due to the mining of the raw material (e.g. coal, gas etc.) were neglected.

Energy lost during these processes are minimal, moreover as they occur in both scenarios (electric and IC) these effects roughly cancel, making it acceptable to leave the effects out.

These efficiencies will impact the energy consumption of an AEV. Therefore, formula (9) is adjusted to entail all efficiency changes. In the worst-case scenario this produces:

$$\text{Power} = \frac{1}{0.319}\Big(0.9\frac{\rho c_d A_{vehicle} v_{non-urban}^3}{2}+0.85*0.95\frac{\rho c_d A_{ve}}{}$$

$$0.85*0.95\frac{m_c v_{urban}^3}{4d}+\frac{m_c v_{non-urban}^3}{4d}+C_{rr}m_c g(v_{non-urban}+0.85$$

Likewise, in the best-case scenario we change formula (9) to:

$$\text{Power} = \frac{1}{0.428}*0.87\Big(0.45\frac{\rho c_d A_{vehicle} v_{non-urban}^3}{2}+0.54*0.9$$

$$0.54*0.9\frac{m_c v_{urban}^3}{4d}+\frac{m_c v_{non-urban}^3}{4d}+C_{rr}m_c g(v_{non-urban}+0.54*$$

Using the average speeds artificially determined through formula (8), which is done in Appendix B & C, the power consumption of AEV can be calculated in a worst (using 10) and best (using 11) case scenario for various braking distances.

Results

With the varying braking distance and ratio of urban an non-urban driving, a range of average speeds was found. Non-realistic average speeds (i.e. speeds far below speed limits in the Netherlands) were filtered from the results, leaving a limited amount of possible braking distances and ratios that could apply to the Netherlands. The ratios 5:1 and 1:5 are therefore excluded from the results entirely. The following graph shows the energy consumption of conventional and AEV for various braking distances and ratios of urban and non-urban driving, if the worst-case scenario is applied.
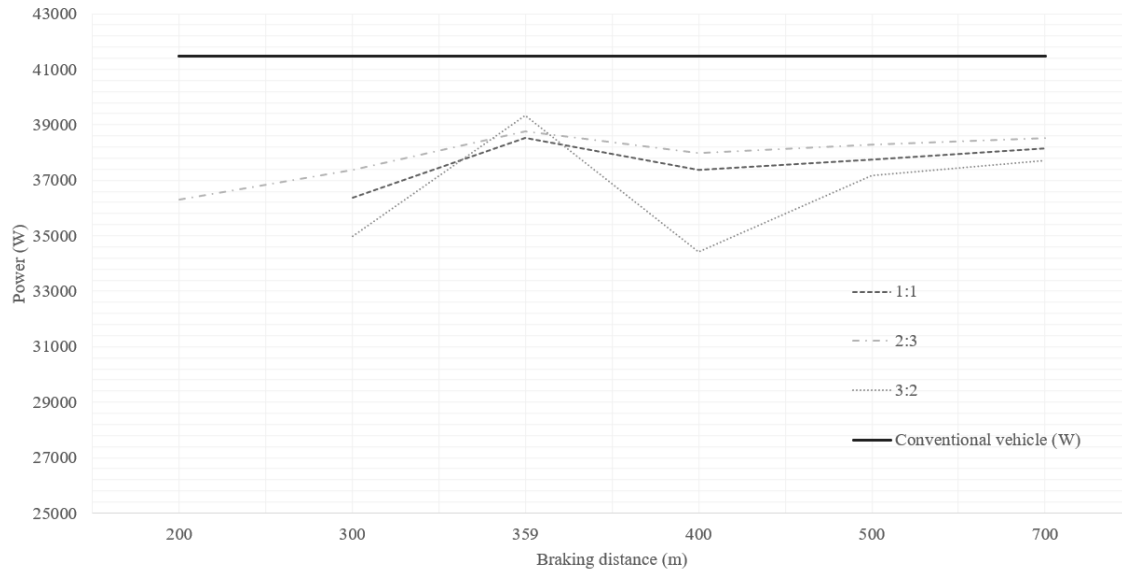
Worst-case Scenario AEV Energy Consumption Compared to Conventional Vehicle Energy Consumption
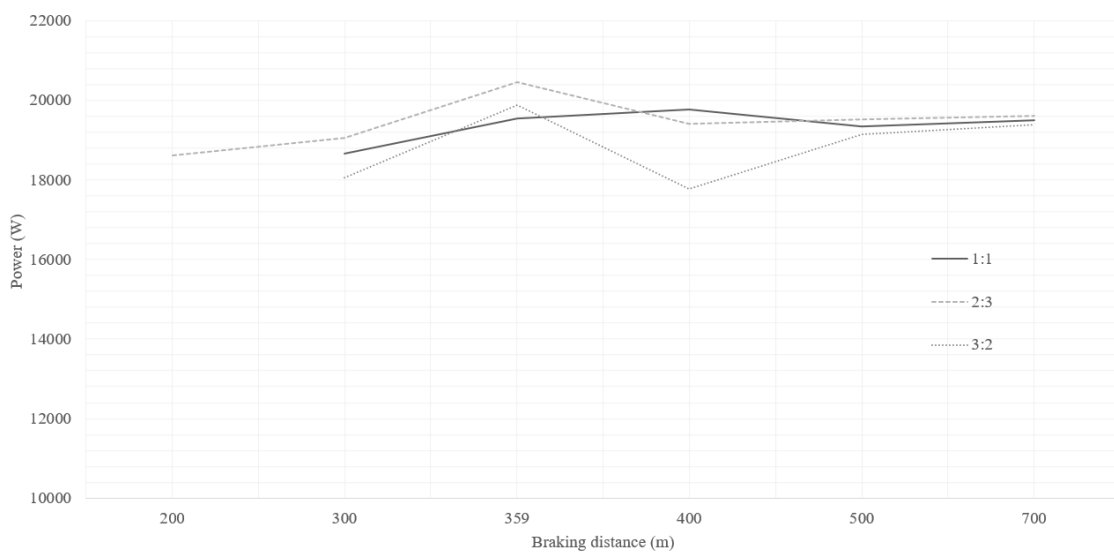
Best-case scenario AEV energy consumption (note that conventional vehicle energy consumption has been left from the graph, as it roughly 20000 Watts higher).

Even in the worst-case scenario, AEV consume slightly less energy than their conventional counterparts. On the other hand, with the best-case scenario energy consumption of AEV drops far below conventional vehicles energy consumption.

However, this shows the energy consumption impact of both the transition to electric *and* autonomous vehicles. If we are to solely look at the impact of a transition to autonomous driving a different result was observed, as now in the worst-

case scenario energy consumption of autonomous vehicles is higher than for that of a conventional vehicle.

Concluding from the figures, a transition to AEV can in both a worst and a best-case scenario only mean a decrease in total energy consumption. In the very worst-case we find a decrease of 5.1%, while a best-case scenario could mean a decrease in energy consumption up to 57.2%. However, re-

sults from the calculations are quite different when the same fuel source is considered. By considering the same fuel source, energy consumption decreases from a transition to electricity as fuel are neglected and thus only the (dis)advantages from autonomous driving become apparent. Looking at the worst-case scenario an increase in energy consumption of 23.2% can be found. Nevertheless, in a best-case scenario, we still find a decrease in energy consumption of 28.2%. Using the results, it can be concluded that just the shift from oil-based fuels to electricity alone provides a decrease in energy consumption of 9.1% and 25.5% in the worst and best-case scenario, respectively. Furthermore, it can be noticed that (especially in the best-case scenarios) energy consumption slightly increases with braking distance, apart from some anomalies that will be discusses below. This increasing energy consumption can be explained by the increasing average speeds in non-urban areas. With increased braking distance, less energy is lost in urban areas; nevertheless, the way the model is oriented, the total energy consumption must remain the same when calculating average speeds. The model then compensates for this decrease in energy consumption by increasing average speed in non-urban areas. Next, when efficiency changes are applied, something interesting happens. Efficiency impacts in urban areas are higher than in non-urban areas: 70% versus 10%, and 119% versus 68%. Therefore, if less energy is consumed in non-urban areas (as breaking distance increases) less energy can be saved and thus the total energy consumption increases. Albeit that in the worst-case scenario this trend cannot be seen clearly in the 3:2 ratio, the other scenarios do follow expectations more evidently, but more on this ratio later.

Additionally, we can observe that in the worst-case scenario, energy consumption in the 1:1 ratio of urban and non-urban driving is consistently lower than energy consumption in the 2:3 ratio. Apart from the anomaly at the 359 meter braking distance, the 3:2 ratio shows similar traits with regards to the 1:1 ratio. This can be explained through the importance of regenerative braking, as in urban driving, more energy can be saved through this method, while in a worst-case scenario, only limited energy savings were deemed possible. In the best-case scenario, we see that the various ratios are more strongly intertwined. This is due to other increases in energy efficiency that were only

considered in the best-case scenario, such as platooning, eco-routing, and eco-driving. This allows for a relatively smaller gap between urban and non-urban efficiency changes. Evidently, these energy efficiency increases (mostly more effective on the highway) can overpower the impact of regenerative breaking. Although, percentage-wise, the efficiency increases in urban areas are higher (119% versus 68%), on average far more energy is consumed on the highway (speed influences energy consumption to the third power) thus allowing this gap to level, as is seen in the best-case scenario.

Lastly, it is remarkable that the energy consumption, if only looked at a transition to AV, is higher for the AV compared to the conventional vehicles (in a worst-case scenario). This can be explained through the energy consumption of the AI controlling the car. Although this AI causes improved driving efficiency (even in a worst-case scenario) it uses around 2000 Watts to accomplish this. Considering that a conventional vehicle only uses 10574 Watts (100% car efficiency), the AI power consumption has a serious impact on the total energy consumption.

Although a conclusion can be drawn from the results, some anomalies do exist. Throughout the results for all ratios in all graphs, a peak in energy consumption can be noted at a braking distance of 359 meters. This peak can be described through the average urban and non-urban speeds associated with it. Looking at the data, we see a recurring theme where (compared to speeds at all other braking distances) urban speed decreases disproportionately while non-urban speeds increase significantly. This causes energy consumption to rise significantly because of the discrepancy between energy efficiency in urban and non-urban areas, as has been discussed earlier. To this list of anomalies the decrease in energy consumption in the ratio 3:2 between the braking distance 300 and 400 meters can be added. Once again, we see the interesting interaction between the energy efficiency differences in urban and non-urban areas. In this case, the average speeds do follow expectations, however, average urban speed decreases more significantly between these two points. This allows for a decrease in energy consumption despite an increase in braking distance.

Conclusion

This paper introduced a model to evaluate the energy consumption of AEV compared to conventional vehicles. A variety of energy efficiency increases in AEV were considered, including eco-routing, platooning and eco-driving. Additionally, the impact on energy consumption due to a transition to electric vehicles was both considered together with and without the autonomous part.

Undergoing a transition to AEV (so both the transition to electric and autonomous vehicles) an 5.1% to 57.2% decrease in energy consumption can be observed for the worst- and best-case respectively. However, a large part of this decrease in energy consumption is due to the electrification of the vehicles. If only the impact of autonomous driving on energy consumption is considered (+23.2% to -28.2%), there exists a serious uncertainty whether the transition is beneficial in terms of the total energy consumption. Both ranges of the impact on energy consumption are wide, therefore further research is needed, helping to map the energy efficiencies of a transition to autonomous driving better. Furthermore, to conclude on a specific figure for the Netherlands further research on braking distances and ratio of urban and non-urban driving is required, as the observed energy consumption differs up to 11.1%, when all energy efficiency changes are kept constant and the scenario is altered.

Discussion

This study has only covered well-known (and documented) areas of car movement where a transition to AEV decreases or increases energy consumption. However, there are many uncertain effects that could have a significant impact on energy efficiency. Firstly, as has been discussed in the research context, a complete transition to autonomous vehicles could decrease the number of accidents by 90%, saving up to 190 billion dollar (Bertoncello and Wee 2015). Although not directly connected to the movement of a car, energy consumption costs regarding damage done to surroundings and vehicles could be seen as part of the total energy consumption of the sector. Therefore, although complicated, the number of accidents that are prevented due to AV could theo-

retically be transcribed into energy savings. Furthermore, a significant decrease in accidents due to AV could make safety measurements in vehicles potentially obsolete. Therefore, it has been suggested that in the future, current safety measures might be removed, decreasing vehicle weight and thus increasing movement efficiency (Meyer et al. 2017). Additionally, electric vehicle weight might decrease due to the further developments in the battery industry. Currently, batteries contribute significantly batteries to the weight of electric vehicles, but with increased use and intensive research batteries might become more powerful, and thus lighter in the future.

Most of these energy savings remain unknown, and could present themselves to be negligible; nevertheless, further research into these topics is needed to conclude their significance. However, some imminant changes in the electricity mix have a far more certain impact on the energy consumption of AEV. Future plans for electricity production in the Netherlands promise a strong increase in renewable energy sources (Schoots et al. 2017). Currently, renewable energy supplies 13.8% of Dutch electricity (CBS 2017e), and the European Union aims to achieve 20% renewable electricty by 2020. Thus, in the short term, renewable energy sources will grow, further advocating for a transition to electric vehicles and possibly AEV. Additionally, we may expect increases in the energy efficiency of the AI driving the car, as the autonomous vehicle market is still developing. In the past, we have seen that AI software updates could increase driving range, suggesting increases in efficiency of the AI.

However, not all future changes have a positive impact on the energy consumption of AEV. As autonomous vehicles do not require any attention or driving skills, various sources suggest possible increases in vehicle use (Meyer et al. 2017)(Wadud et al. 2016). An increase in use of passenger vehicles will bring along an increase in the total energy consumption of the passenger vehicles transport sector, which could potentially diminish any energy saved due to the transition to AEV.

Although one would expect average speed in urban areas to go up as braking distance increases (less braking incidents meaning less slowing down, thus keeping a higher average speed), some of the results, as has been discussed, show different behaviour. Therefore, although the model suggests that energy consumption drops when braking dis-

tance is at 400 meters, it could be a flaw in the model. Similarly the peak experienced throughout the ratios at a braking distance of 359 meters might be due to a flaw in the model. A potential problem could be that the model only considers two areas of driving combined with two different braking distances, while in reality, many more areas of driving exist. Therefore, it is suggested that future models differentiate more areas of driving while simultaneously acknowledging the various braking distances in these areas. However, this was beyond the possibilities of this study due to a lack of data. Nevertheless, this study believes that creating a more extensive model will smooth out the aforementioned inconsistencies.

Taking the variety of uncertainties (and unknowns) in energy efficiency impacts of autonomous driving into consideration, the effect of a transition to AEV is still difficult to quantify. This can clearly be seen in the results of this study, as their range is extremely wide. Therefore, further research on these energy efficiency changes is desperately needed, as car manufacturers have predicted that production of fully autonomous cars could start within five years. To find a more accurate number for the Netherlands, further research on braking distance in urban and non-urban areas as well as an accurate ratio for urban and non-urban driving will be needed. Luckily, CBS is currently looking into this topic, and will be able to provide figures by the start of 2019.

Works Cited

Ambel, CC. 2017. Roadmap to climate-friendly land freight and buses in Europe. Technical Report March. Transport and Environment.

Bashmakov, IA, Mulugetta, Y, Bruckner, T. 2014. Energy Systems. Technical report. Intergovernmental Panel on Climate Change.

Bertoncello, M, Wee, D. 2015. Ten ways autonomous driving could redefine the automotive world. Mckinsey and Company.

Boriboonsomsin, K, Barth, MJ, Member, S, Zhu, W, Vu, A. 2012. Eco-Routing Navigation System Based on Multisource Historical and Real-Time Traffic Information. IEEE Transactions On Intelligent Transportation Systems. 13(4):1694–1704.

CBS. 2016a. Aantal personenauto's in Nederland 2016. Technical report. Centraal Bureau voor de Statistiek.

CBS. 2016b. Elektriciteit en warmte; productie en inzet naar energiedrager. Technical report. Centraal Bureau voor de Statistiek.

CBS. 2016c. Personenmobiliteit in Nederland; persoonskenmerken en reismotieven, 2016. Technical report. Centraal Bureau voor de Statistiek.

CBS. 2017a. Aantal personenauto's in Nederland 2018. Technical report. Centraal Bureau voor de Statistiek.

CBS. 2017b. Bevolkingaantal 2016. Technical report. Centraal Bureau voor de Statistiek.

CBS. 2017c. Elektriciteit en warmte; productie en inzet naar energiedrager. Technical report. Centraal Bureau voor de Statistiek.

CBS. 2017d. Energieverbruik door verkeer en vervoer, 1990-2015. Technical report. Centraal Bureau voor de Statistiek.

CBS. 2017e. Hernieuwbare elektriciteit; productie en vermogen. Technical report. Centraal Bureau voor de Statistiek.

Conti, JJ, Holdtenburg, PD, Diefenderfer, JR. 2016. Transportation Sector Energy Consumption. Technical report. US Energy Information Administration.

Duan, K, McDaniel, C, Muller, A, Yokoto, B. 2007. Effects of Highway Slipstreaming on California Gas Consumption.

EIA. 2018. Monthly Energy Review April 2018. Technical report. US Energy Information Administration.

Gillespie, D. 1992. Fundamentals of Vehicle Dynam- ics. Society of Automotive Engineers, Inc. doi: 10.4271/R-114. NIHMS150003.

He, Y, Rios, J, Chowdhury, M, Pisu, P, Bhavsar, P. 2012. Forward power-train energy management modeling for assessing benefits of integrating predictive traffic data into plug-in-hybrid elec- tric vehicles. Transportation Research Part D: Transport and Environment. 17(3):201–207. doi: 10.1016/j.trd.2011.11.001.

Herzog, T, Baumert, K, Pershing, J. 2005. Navigating the Numbers. World Resources Institute. Hills, P. 1996. What is induced traffic? Transportation. 23(1):5–16. doi: 10.1007/BF00166216.

Mackay, DJC. 2009. Sustainable Energy — with- out the hot air. 2. UIT Cambridge Ltd. doi: 10.1109/PES.2004.1373296.

Mensing, F, Bideaux, E, Trigui, R, Tattegrain, H. 2013. Trajectory optimization for eco-driving taking into account traffic constraints. Transportation Research Part D: Transport and Environment. 18(1):55–61. doi: 10.1016/j.trd.2012.10.003.

Meyer, J, Becker, H, Bösch, PM, Axhausen, KW. 2017. Autonomous vehicles: The next jump in accessibilities? Research in Transportation Economics. 62:80–91. doi: 10.1016/j.retrec.2017. 03.005.

NHTSA. 2008. National Motor Vehicle Crash Causation Survey Report to Congress. Technical Report July. U.S. Department of Transportation.

Pachauri, RK, Meyer, L, Van Ypersele, JP, Brinkman, S, Van Kesteren, L, Leprince-Ringuet, N, Van Boxmeer, F. 2014. Climate Change 2014 Synthe- sis Report. Technical report. Intergovernmental Panel for Climate Change.

Paling, J. 2013. Omzetting primaire energiedragers in elektriciteit. Technical report. TU Delft.

Schoots, K, Hekkenberg, D, Hammingh, P. 2017. Na- tionale Energieverkenning 2017. ECN-O–14-036. p. 1–276. doi: ECN-O–16-035.

Société de Technologie Michelin. 2003. The tyre, rolling resistance and fuel savings. Technical re- port. Société de Technologie Michelin.

Wadud, Z, MacKenzie, D, Leiby, P. 2016. Help or hindrance? The travel, energy and carbon impacts of highly automated vehicles. Transportation Re- search Part A: Policy and Practice. 86:1–18. doi: 10.1016/j.tra.2015.12.001.

Welling, M. 2017. Intelligentie per Kilowattuur. Het Financieele Dagblad, [date accessed 2018 Mar 20].

Yazdanie, M, Noembrini, F, Dossetto, L, Boulouchos, K. 2014. A comparative analysis of well-to- wheel primary energy demand and greenhouse gas emissions for the operation of alternative and conventional vehicles in Switzerland, considering various energy carrier production pathways. Journal of Power Sources. 249:333–348. doi: 10.1016/j.jpowsour.2013.10.043.

Zabat, M, Stabile, N, Farascaroli, S, Browand, F. 1995. The Aerodynamic Performance Of Pla- toons: A Final Report. California Partners for Ad- vanced Transit and Highways (PATH).

Zhu, H, Yang, Z. 2011. Simulation of the aerodynamic interaction of two generic sedans moving very closely. 2011 International Conference on Electric Information and Control Engineering, ICE- ICE 2011 - Proceedings. p. 2595–2600. doi: 10.1109/ICE-ICE.2011.5777136.

A Model Values Calculations

A.1 An Average Car According to Mackay

The values used by Mackay to calculate energy consumption of an average car are given here:

Table 2: Vehicle Model Calculation Parameters (Mackay)

Parameter	Value
Air density	1.3 kg/m^3
Vehicle Area	3 m^2
Mass of car (conventional)	1000 kg
Rolling resistance	0.010
Gravitation of earth	9.81 m/s^2
Drag coefficient car	0.33

A.2   Average Power Consumption Dutch passenger vehicles

The following facts, provided by CBS, make it possible to calculate the average power consumption of a single passenger vehicle in the Netherlands.

Data type	Amount
Number of Dutch passenger vehicles (CBS 2016a)	8100864
Total Dutch passenger vehicles energy consumption (CBS 2017d)	259 PJ
Average car travel time per person per day (CBS 2016c)	19.76 minutes
Number of inhabitants of the Netherlands (**?**)CBS 2017b)	169790120
Average efficiency of conventional car engine (Ambel 2017)	30 %

First it is necessary to calculate average travel time per car per year. This is done through:

$$\text{Yearly travel time per person} * \frac{\text{\# of inhabitants}}{\text{\# of passenger vehicles}} =$$

$$60 * 365 * 19.76 * \frac{169790120}{8100864} = 906879 \text{ Seconds per car.} \quad (12)$$

Then we simply divide total energy consumption by this number and keeping car engine efficiency in mind:

$$\frac{259 * 10^{15}}{906879 * 8100864} = 35249 \text{ Watts per car} \Rightarrow 35249 * 0.3 = 10574.8 \text{ Watts per car for car movement.} \quad (13)$$

Do note that this energy consumption is only from tank to wheel. Total energy consumption should be determined through the Well-to-Wheel (WTW) analysis. Especially during the refining of oil energy losses occur. These losses are well known and average around 15% (Pachauri et al. 2014). Therefore, total energy consumption for car movement is 41469 Watts ($35249 * \frac{1}{0.85}$).

A.3   Average Weight Dutch Passenger Vehicles

CBS also gathers information on the amount of passenger vehicles that fit in a certain weight class (CPBS 2017a). The data is slightly more recent than used in A.1, however car weight is not strongly volatile, so this causes no problem.

$$\frac{\text{Total Weight Dutch Passenger Vehicles}}{\text{Amount of Dutch passenger vehicles}} = \frac{9748327375}{8373244} = 1164.22 \text{ kg} \quad (14)$$

A.4   Average distance covered per car per day

CBS also has data on the average travel distance by passenger vehicle per person (CBS 2016c). To translate this to the average travel distance per car we write, similarly to 12,

$$\text{Average travel distance per person} * \frac{\text{\# of inhabitants}}{\text{\# of passenger vehicles}} = 15.14 * \frac{169790120}{8100864} = 31.73 \text{ km.} \quad (15)$$

A.5   Average Travel Time per Day per Car

Simply using the total yearly travel time per car found in A.1 we find

$$\frac{\text{Yearly travel time per car}}{\text{\# of days in the year}} = \frac{906879}{365} = 2484.6 \text{ s.} \quad (16)$$

A.6   Efficiency of Electric Vehicle Engine

To calculate the efficiency of an electric engine we need to consider production efficiency of Dutch electricity as well as the energy efficiency of the engine itself. Ambel states that the average electric engine efficiency lies at 90%. However, energy losses (5%) due to inversions from AC to DC and vice versa and battery charging inefficiencies (5%) have to be considered. Furthermore, data on the Dutch electricity can be found through CBS (CBS 2016b), while average power plant efficiencies were determined in Paling (2013). The results of these inefficiencies are combined in the following table.

Table 3: Electricity Production Efficiency

	% of total electicity production	Average power plant efficiency (%)	
Natural Gas	0.42	45-60	
Coal	0.36	30-45	
Fuel oil	0.00087	30-45	
Other fossil fuels	0.036	30-45	Worst-case efficiency
Solar	0.010	100	31.9 %
Wind	0.069	100	
Hydro	0.00085	100	
Biomass	0.045	20-40	
Nuclear	0.037	30-35	Best-case efficiency
Other energy carriers	0.026	30-45	42.8%

B   Calculations with the Vehicle Model

In this section all calculations with the vehicle model can be found. The values found from the calculations were used to set up graphs in Wolfram Mathematica, which were in turn used to determine average speeds for urban and non-urban areas. The model formula has been rewritten as

$$\text{Power} = A_1 v_{non-urban}^3 + A_1 v_{urban}^3 + B_1 v_{urban}^3 + B_2 v_{non-urban}^3 + C_1(v_{urban} + v_{non-urban}), \quad (17)$$

where A_1, $B_{1/2}$ and C_1 are calculated using values given in the table below (in combination with a varying urban braking distance and ratio urban/non-urban driving), which have already been backed scientifically in the research context, methodology and appendices. Additionally when calculating energy consumption of AEV the formula is adjusted for the increased weight. The altered value is given below and is based on the increase in weight of vehicles on the market that have both an electric and IC version.

Using these values in combination with formula (17), which is rewritten to:

$$\text{Power} = (A_1 + B_1)(\frac{x_{urban}}{t_{urban}})^3 + C_1(\frac{x_{urban}}{t_{urban}}) +$$

$$C_1(\frac{x_{non-urban}}{t_{non-urban}}) + (A_1 + B_2)(\frac{x_{non-urban}}{t_{non-urban}})^3 +, \quad (18)$$

We can find a range of solutions of the average time spent in both urban and non-urban areas. These solutions are found by solving the following equations, which obey specific parameters described in the table.

Table 4: Vehicle Model Calculation Parameters

Parameter	Value
Braking distance	100 to 700 m
Air density	1.2466 kg/m^3
Vehicle Area	3 m^2
Mass of car (conventional)	1164 kg
Mass of car (electric)	1414 kg
Rolling resistance	0.012
Gravitation of earth	9.81 m/s^2
Drag coefficient car	0.33
Braking distance non-urban areas(m)	Non-urban driving distance
Distance covered per car per day	31.73 km
Average power consumption of a Dutch passenger vehicle	10574

C   Calculating the average speeds

Solving the formulas described in Appendix B with the boundary condition that $t_{urban} + t_{non-urban}$ is equal to 2484.6 seconds the following results were found (remember that the ratios are connected to specific values, thus to get the average speed one simply has to divide this distance by the time).

D   Tables for the various scenarios

In this appendix the tables on which the graphs in the results are based can be found. Additionally these tables were used to calculate the maximum de- and increase of energy consumption.

Table 5: Urban braking distance 100 meter

Urban	26444.0802
Non-urban	5288.816041
Solve	$6.437(\frac{26444.0802}{x})^3 + 137.02(\frac{26444.0802}{x}) + 137.02(\frac{5288.81}{y}) + 0.727110533(\frac{5288.816}{y})^3 = 10574$

Urban	5288.816041
Non-urban	26444.0802
Solve	$6.437(\frac{5288.816041}{x})^3 + 137.02(\frac{5288.816041}{x}) + 137.02(\frac{26444.0802}{y}) + 0.639075707(\frac{26444.0802}{y})^3 = 10574$

Urban	15866.44812
Non-urban	15866.44812
Solve	$6.437(\frac{15866.44812}{x})^3 + 137.02(\frac{15866.44812}{x}) + 137.02(\frac{15866.44812}{y}) + 0.653748178(\frac{15866.44812}{y})^3 = 10574$

Urban	12693.1585
Non-urban	19039.73775
Solve	$6.437(\frac{12693.1585}{x})^3 + 137.02(\frac{12693.1585}{x}) + 137.02(\frac{19039.73775}{y}) + 0.647634648(\frac{19039.73775}{y})^3 = 10574$

Urban	19039.73775
Non-urban	12693.1585
Solve	$6.437(\frac{19039.73775}{x})^3 + 137.02(\frac{19039.73775}{x}) + 137.02(\frac{12693.1585}{y}) + 0.662918472(\frac{12693.1585}{y})^3 = 10574$

Table 6: Urban braking distance 200 meter

Urban	26444.0802
Non-urban	5288.816041
Solve	$3.527(\frac{26444.0802}{x})^3 + 137.02(\frac{26444.0802}{x}) + 137.02(\frac{5288.81}{y}) + 0.727110533(\frac{5288.816}{y})^3 = 10574$
Urban	5288.816041
Non-urban	26444.0802
Solve	$3.527(\frac{5288.816041}{x})^3 + 137.02(\frac{5288.816041}{x}) + 137.02(\frac{26444.0802}{y}) + 0.639075707(\frac{26444.0802}{y})^3 = 10574$
Urban	15866.44812
Non-urban	15866.44812
Solve	$3.527(\frac{15866.44812}{x})^3 + 137.02(\frac{15866.44812}{x}) + 137.02(\frac{15866.44812}{y}) + 0.653748178(\frac{15866.44812}{y})^3 = 10574$
Urban	12693.1585
Non-urban	19039.73775
Solve	$3.527(\frac{12693.1585}{x})^3 + 137.02(\frac{12693.1585}{x}) + 137.02(\frac{19039.73775}{y}) + 0.647634648(\frac{19039.73775}{y})^3 = 10574$
Urban	19039.73775
Non-urban	12693.1585
Solve	$3.527(\frac{19039.73775}{x})^3 + 137.02(\frac{19039.73775}{x}) + 137.02(\frac{12693.1585}{y}) + 0.662918472(\frac{12693.1585}{y})^3 = 10574$

Table 7: Urban braking distance 359 meter

Urban	26444.0802
Non-urban	5288.816041
Solve	$2.238(\frac{26444.0802}{x})^3 + 137.02(\frac{26444.0802}{x}) + 137.02(\frac{5288.81}{y}) + 0.727110533(\frac{5288.816}{y})^3 = 10574$
Urban	5288.816041
Non-urban	26444.0802
Solve	$2.238(\frac{5288.816041}{x})^3 + 137.02(\frac{5288.816041}{x}) + 137.02(\frac{26444.0802}{y}) + 0.639075707(\frac{26444.0802}{y})^3 = 10574$
Urban	15866.44812
Non-urban	15866.44812
Solve	$2.238(\frac{15866.44812}{x})^3 + 137.02(\frac{15866.44812}{x}) + 137.02(\frac{15866.44812}{y}) + 0.653748178(\frac{15866.44812}{y})^3 = 10574$
Urban	12693.1585
Non-urban	19039.73775
Solve	$2.238(\frac{12693.1585}{x})^3 + 137.02(\frac{12693.1585}{x}) + 137.02(\frac{19039.73775}{y}) + 0.647634648(\frac{19039.73775}{y})^3 = 10574$
Urban	19039.73775
Non-urban	12693.1585
Solve	$2.238(\frac{19039.73775}{x})^3 + 137.02(\frac{19039.73775}{x}) + 137.02(\frac{12693.1585}{y}) + 0.662918472(\frac{12693.1585}{y})^3 = 10574$

Table 8: Urban braking distance 400 meter

Urban	26444.0802
Non-urban	5288.816041
Solve	$2.072(\frac{26444.0802}{x})^3 + 137.02(\frac{26444.0802}{x}) + 137.02(\frac{5288.81}{y}) + 0.727110533(\frac{5288.816}{y})^3 = 10574$
Urban	5288.816041
Non-urban	26444.0802
Solve	$2.072(\frac{5288.816041}{x})^3 + 137.02(\frac{5288.816041}{x}) + 137.02(\frac{26444.0802}{y}) + 0.639075707(\frac{26444.0802}{y})^3 = 10574$
Urban	15866.44812
Non-urban	15866.44812
Solve	$2.072(\frac{15866.44812}{x})^3 + 137.02(\frac{15866.44812}{x}) + 137.02(\frac{15866.44812}{y}) + 0.653748178(\frac{15866.44812}{y})^3 = 10574$
Urban	12693.1585
Non-urban	19039.73775
Solve	$2.072(\frac{12693.1585}{x})^3 + 137.02(\frac{12693.1585}{x}) + 137.02(\frac{19039.73775}{y}) + 0.647634648(\frac{19039.73775}{y})^3 = 10574$
Urban	19039.73775
Non-urban	12693.1585
Solve	$2.072(\frac{19039.73775}{x})^3 + 137.02(\frac{19039.73775}{x}) + 137.02(\frac{12693.1585}{y}) + 0.662918472(\frac{12693.1585}{y})^3 = 10574$

Table 9: Urban braking distance 500 meter

Urban	26444.0802
Non-urban	5288.816041
Solve	$1.781(\frac{26444.0802}{x})^3 + 137.02(\frac{26444.0802}{x}) + 137.02(\frac{5288.81}{y}) + 0.727110533(\frac{5288.816}{y})^3 = 10574$
Urban	5288.816041
Non-urban	26444.0802
Solve	$1.781(\frac{5288.816041}{x})^3 + 137.02(\frac{5288.816041}{x}) + 137.02(\frac{26444.0802}{y}) + 0.639075707(\frac{26444.0802}{y})^3 = 10574$
Urban	15866.44812
Non-urban	15866.44812
Solve	$1.781(\frac{15866.44812}{x})^3 + 137.02(\frac{15866.44812}{x}) + 137.02(\frac{15866.44812}{y}) + 0.653748178(\frac{15866.44812}{y})^3 = 10574$
Urban	12693.1585
Non-urban	19039.73775
Solve	$1.781(\frac{12693.1585}{x})^3 + 137.02(\frac{12693.1585}{x}) + 137.02(\frac{19039.73775}{y}) + 0.647634648(\frac{19039.73775}{y})^3 = 10574$
Urban	19039.73775
Non-urban	12693.1585
Solve	$1.781(\frac{19039.73775}{x})^3 + 137.02(\frac{19039.73775}{x}) + 137.02(\frac{12693.1585}{y}) + 0.662918472(\frac{12693.1585}{y})^3 = 10574$

Table 10: Urban braking distance 700 meter

Urban	26444.0802
Non-urban	5288.816041
Solve	$1.448(\frac{26444.0802}{x})^3 + 137.02(\frac{26444.0802}{x}) + 137.02(\frac{5288.81}{y}) + 0.727110533(\frac{5288.816}{y})^3 = 10574$
Urban	5288.816041
Non-urban	26444.0802
Solve	$1.448(\frac{5288.816041}{x})^3 + 137.02(\frac{5288.816041}{x}) + 137.02(\frac{26444.0802}{y}) + 0.639075707(\frac{26444.0802}{y})^3 = 10574$
Urban	15866.44812
Non-urban	15866.44812
Solve	$1.448(\frac{15866.44812}{x})^3 + 137.02(\frac{15866.44812}{x}) + 137.02(\frac{15866.44812}{y}) + 0.653748178(\frac{15866.44812}{y})^3 = 10574$
Urban	12693.1585
Non-urban	19039.73775
Solve	$1.448(\frac{12693.1585}{x})^3 + 137.02(\frac{12693.1585}{x}) + 137.02(\frac{19039.73775}{y}) + 0.647634648(\frac{19039.73775}{y})^3 = 10574$
Urban	19039.73775
Non-urban	12693.1585
Solve	$1.448(\frac{19039.73775}{x})^3 + 137.02(\frac{19039.73775}{x}) + 137.02(\frac{12693.1585}{y}) + 0.662918472(\frac{12693.1585}{y})^3 = 10574$

Table 11: Urban braking distance 100 meters

100 meter Ratio	Urban (s)	N. S. = No Solution Non-urban (s)	Total	Urban (km/h)	Non-urban (km/h)
5:01	N. S.	N. S.	0	N. S.	N. S.
1:05	1271	1213	2484	15.0	78.5
1:01	N. S.	N. S.	0	N. S.	N. S.
2:03	N. S.	N. S.	0	N. S.	N. S.
3:02	N. S.	N. S.	0	N. S.	N. S.

Table 12: Urban braking distance 200 meters

200 meter Ratio	Urban (s)	N. S. = No Solution Non-urban (s)	Total	Urban (km/h)	Non-urban (km/h)
5:01	N. S.	N. S.	0	N. S.	N. S.
1:05	1287	1202	2489	14.8	79.2
1:01	1524	965	2489	37.5	59.2
2:03	1496	991	2487	30.5	69.2
3:02	N. S.	N. S.	0	N. S.	N. S.

Table 13: Urban braking distance 300 meters

300 meter Ratio	Urban (s)	Non-urban (s)	Total	Urban (km/h)	Non-urban (km/h)
5:01	2103	383.7	2486.7	45.3	49.6
1:05	1284	1200	2484	14.8	79.3
1:01	1642	842	2484	34.8	67.8
2:03	1536	948	2484	29.7	72.3
3:02	1745	739	2484	39.3	61.8

Table 14: Urban braking distance 359 meters

359 meter Ratio	Urban (s)	Non-urban (s)	Total	Urban (km/h)	Non-urban (km/h)
5:01	2141	344	2485	44.5	55.3
1:05	1266	1209	2475	15.0	78.7
1:01	1691	793	2484	33.8	72.0
3:02	1839	645	2484	37.3	70.8
2:03	1600	884	2484	28.6	77.5

Table 15: Urban braking distance 400 meters

400 meter Ratio	Urban (s)	Non-urban (s)	Total	Urban (km/h)	Non-urban (km/h)
5:01	2142	341	2483	44.4	55.8
1:05	1288	1197	2485	14.8	79.5
1:01	1669	815	2484	34.2	70.1
2:03	1551	933	2484	29.5	73.5
3:02	1794	690	2484	38.2	66.2

Table 16: Urban braking distance 500 meters

500 meter Ratio	Urban (s)	Non-urban (s)	Total	Urban (km/h)	Non-urban (km/h)
5:01	2119	329.9	2448.9	44.9	57.7
1:05	1287	1201	2488	14.8	79.3
1:01	1683	801	2484	33.9	71.3
2:03	1561	923	2484	29.3	74.3
3:02	1813	671	2484	37.8	68.1

Table 17: Urban braking distance 700 meters

700 meter Ratio	Urban (s)	Non-urban (s)	Total	Urban (km/h)	Non-urban (km/h)
5:01	2154	305	2459	44.2	62.4
1:05	1216	1258	2474	15.7	75.7
1:01	1697	787	2484	33.7	72.6
2:03	1570	914	2484	29.1	75.0
3:02	1831	653	2484	37.4	70.0

Table 18: AEV and conventional energy consumption after efficiency changes (worst-case scenario)

	1:1 (W)	2:3 (W)	3:2 (W)	Conventional vehicle (W)
200 (m)		36316		41496
300 (m)	36388	37372	34997	41469
359 (m)	38514	38772	39333	41469
400 (m)	37368	37985	34417	41469
500 (m)	37746	38285	37164	41469
700 (m)	38149	38525	35801	41469

Table 19: AEV energy consumption after efficiency changes (best-case scenario)

	1:1 (W)	2:3 (W)	3:2 (W)
200 (m)		18619	
300 (m)	18658	19055	18060
359 (m)	19533	20446	19887
400 (m)	19769	19411	17768
500 (m)	19351	19526	19151
700 (m)	19506	19616	18486

Table 20: Sole impact of autonomous driving on energy consumption rate

	Worst-case = w-c			Best-case = b-c		
	1:1 w-c (W)	2:3 w-c (W)	3:2 w-c (W)	1:1 b-c (W)	2:3 b-c (W)	3:2 b-c (W)
200 (m)		11585			7404	
300 (m)	11608	11921	11164	7839	8006	7588
359 (m)	12286	13031	12547	8207	8590	8355
400 (m)	11920	12117	11567	8306	8156	7824
500 (m)	12041	12213	11856	8130	8204	8047
700 (m)	12169	12289	12033	8196	8242	8140